

Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Blaž Meden

**Sistem za sledenje znanim objektom v
spremenljivih svetlobnih pogojih na sceni s
pomočjo posnetkov IR in vidnega spektra**

MAGISTRSKO DELO
MAGISTRSKI PROGRAM DRUGE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

izr. prof. dr. Peter Peer
MENTOR

Ljubljana, 2016

© 2016, Blaž Meden

Rezultati magistrskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov magistrskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

IZJAVA O AVTORSTVU MAGISTRSKEGA DELA

Spodaj podpisani izjavljam, da sem avtor dela, da slednje ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo, za pridobitev naziva na univerzi ali drugem visokošolskem zavodu, razen v primerih kjer so navedeni viri.

S svojim podpisom zagotavljam, da:

- sem delo izdelal samostojno pod mentorstvomizr. prof. dr. Petra Peera,
- so elektronska oblika dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko in
- soglašam z javno objavo elektronske oblike dela v zbirki “Dela FRI”.

— Blaž Meden, Ljubljana, maj 2016.

Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Blaž Meden

Sistem za sledenje znanim objektom v spremenljivih svetlobnih pogojih na sceni s pomočjo posnetkov IR in vidnega spektra

POVZETEK

V tem raziskovalnem delu naslovimo problematiko sledenja večih objektov pod vplivom negotovih svetlobnih razmer. Da bi rešili problem dinamične osvetlitve implementiramo dva dobro znana sledilna pristopa, ki operirata v dveh ločenih domenah. Pri tem se prvi sledilnik zanaša na aktivni sistem oznak, ki deluje v bližnjem infrardečem spektru, drugi pa izkorišča najsodobnejši obstoječi detektor oseb in barvno informacijo opazovane scene v vidnem spektru. Namesto uporabe le enega od predstavljenih sledilnikov izboljšamo zmogljivost sledenja objektov v zaprtih prostorih z združitvijo obeh pristopov. To hkrati pomeni, da je implementiran sistem primeren za uporabo v povsem praktičnih aplikacijah, saj je načrtovan tako, da podpira sledenje do 16 gibajočih se objektov. V prvem delu naloge predstavimo obstoječe raziskovalno delo, ki je bilo opravljeno na tem področju, zatem opišemo podrobnosti implementacije in vrednotenja sistema ter na koncu predlagamo tudi množico direktiv za nadaljnje raziskovalno delo in možne izboljšave našega sistema.

Ključne besede: računalniški vid, sledenje objektov, Kalmanov filter, filter z delci, Madžarska metoda, infrardeče oznake

University of Ljubljana
Faculty of Computer and Information Science

Blaž Meden

Tracking system of known objects in unstable lightning conditions on stage using IR and visible spectrum video inputs

ABSTRACT

In this research work we address the multiple object tracking problem under challenging illumination conditions. To solve the changing illumination problem, we implement two well known tracking approaches, which are operating in two separated domains. One tracker relies on active marker system, which operates in near infrared spectrum and the other one exploits existing state of the art pedestrian detector and color information by observing the scenery in visible spectrum. By integrating these two approaches, we intend to improve the performance of object tracking in closed areas, rather than using only one of the introduced trackers. This also means that the implemented system is usable in many practical applications, as it is designed to track up to 16 moving objects. In the first part of the thesis, we cover existing research work, which was done in this field, then we describe implementation and evaluation details of the system and finally we propose a set of directions for future research and possible improvements of our system.

Key words: computer vision, object tracking, Kalman filter, particle filter, Hungarian algorithm, infrared markers

ZAHVALA

V prvi vrsti bi se zahvalil mentorju izr. prof. dr. Petru Peeru za vso pomoč, potrpežljivost in spodbudo, zatem as. Bojanu Klemencu za vso dodatno pomoč in viš. pred. dr. Francu Policardiju za vse konstruktivne ideje, ter tudi prijatelju Tinetu Misleju, za prisotnost in pomoč pri zajemanju testnih posnetkov. Vsekakor pa gre zahvala tudi staršem za vso podporo med študijem in še posebej med izdelavo te magistrske naloge.

— Blaž Meden, Ljubljana, maj 2016.

SEZNAM KRATIC

- DBT - Detection Based Tracking - sledenje na bazi detekcije objektov
- DFT - Detection Free Tracking - sledenje brez detekcije objektov
- GNN - Global Nearest Neighbour - iskanje najbližjih globalnih sosedov
- MAP - maximum a posteriori probability - maksimalna pogojna verjetnost
- KLT - Kanade-Lucas-Tomasi - kratica avtorjev Kanade, Lucas in Tomasi
- HOG - Histograms of Oriented Gradients - histogrami usmerjenih gradientov
- FFT - Fast Fourier Transform - hitra Fourierova transformacija
- POM - Probabilistic Occupancy Map - zemljevid verjetnostne zasedenosti
- NCC - Normalized Cross Correlation - normalizirana križna korelacija
- RGB - red green blue - barvni kanali – rdeča, zelena, modra
- HSV - hue saturation value - barvni kanali – odtenek, nasičenost, intenziteta
- SIFT - Scale-Invariant Feature Transform - skalarno invariantne lokalne značilke
- SPM - Spatial Pyramid Matching - prostorsko piramidno ujemanje
- SVM - Support Vector Machine - klasifikacija s podpornimi vektorji
- PCA - Principal Component Analysis - analiza glavnih komponent
- LPB - Local Binary Patterns - lokalni binarni vzorci
- NCV - Near Constant Velocity model - model gibanja
- CRF - Conditional Random Fields - pogojno naključna polja
- FF - Floor Field - talno polje
- HMM - Hidden Markov Model - skriti Markov model
- CTM - Correlated Topic Model - koreliran tematski model
- DPMM - Dirichlet Process Mixture Model - Dirichletov model mešanih procesov
- EOM - Explicit Occlusion Model - eksplicitni model prekrivanja
- EKF - Extended Kalman filter - razširitev Kalmanovega filtra
- UKF - Unscented Kalman filter - razširitev Kalmanovega filtra
- JPDA - Joint Probabilistic Data Association - verjetnostno združena podatkovna asociacija
- MHT - Multiple Hypothesis Tracking - sledenje večim hipotezam
- MWIS - Maximum-Weight Independent Set - neodvisna množica z maksimalno težo
- MOTA - Multiple Object Tracking Accuracy - metrika vrednotenja točnosti sledenja
- TDE - Tracking Distance Error - metrika vrednotenja natančnosti sledenja
- MOTP - Multiple Object Tracking Precision - metrika vrednotenja natančnosti sledenja
- FP - false positive - lažna detekcija, FN - false negative - zgrešena detekcija
- MME - mismatch event - dogodek zamenjave identitete
- GT - ground truth - zlati standard
- MT - mostly tracked - večinsko pokrito, PT - partly tracked - delno pokrito
- ML - mostly lost - večinsko izgubljeno, FM - fragmentation - fragmentacija

KAZALO

Povzetek	i
Abstract	iii
Zahvala	v
Seznam kratic	vii
1 Uvod	1
1.1 Motivacija	1
1.2 Prispevek naloge	2
1.3 Metodologija	2
1.4 Pregled naloge	3
1.5 Terminologija	3
2 Formuliranje problema	5
3 Kategorizacija metod	7
3.1 Glede na inicializacijo	7
3.1.1 Sledenje z detekcijo objektov	7
3.1.2 Sledenje brez predhodne detekcije	8
3.2 Glede na način procesiranja	8
3.2.1 Sprotno procesiranje	9
3.2.2 Paketno procesiranje	9
3.3 Glede na matematično ozadje	9
3.3.1 Verjetnostne metode	10
3.3.2 Deterministične metode	10

3.4	Glede na področje uporabe	10
3.4.1	Športni scenariji	10
3.4.2	Letalstvo in scenariji iz zraka	11
3.4.3	Sledenje pešcev in prometa	11
3.4.4	Splošna uporaba	12
4	Komponente sledenja	13
4.1	Vizualni modeli	13
4.1.1	Vrste značilk	14
4.1.2	Enostavni modeli	16
4.1.3	Sestavljeni modeli	19
4.2	Modeli gibanja	20
4.2.1	Linearni modeli	21
4.2.2	Nelinearni modeli	23
4.3	Modeli interakcije	24
4.3.1	Socialni modeli	24
4.3.2	Modeli gibanja množic	26
4.4	Modeli izključevanja	27
4.4.1	Izključevanje na nivoju detekcije	28
4.4.2	Izključevanje na nivoju trajektorij	28
4.5	Reševanje prekrivanja	29
4.5.1	Del proti celoti	29
4.5.2	Testiranje hipotez	30
4.5.3	Predpomnjenje in obnovitev	31
4.5.4	Ostalo	32
4.6	Primeri verjetnostnih metod	32
4.6.1	Rekurzivni Bayesovi filtri	32
4.6.2	Filtri z delci	33
4.6.3	Filtri JPDA	33
4.6.4	Metoda MHT	33
4.7	Primeri determinističnih optimizacij	34
4.7.1	Ujemanje dvodelnih grafov	34
4.7.2	Dinamično programiranje	34

4.7.3	Maksimalno pretočno omrežje z minimalno ceno	35
4.7.4	Pogojno naključna polja	35
4.7.5	Neodvisna množica z maksimalno težo	35
5	Praktični del	37
5.1	Študija primerov uporabe	37
5.1.1	Primeri uporabe svetlobnih oznak	37
5.1.2	Primeri možne osvetlitve prizorišča	38
5.1.3	Primeri možne postavitve kamer	39
5.2	Strojna oprema	39
5.2.1	Kamere	39
5.2.2	Oznake IR	40
5.3	Programski del	41
5.3.1	Sledilnik v spektru IR	41
5.3.2	Sledilnik v vidnem spektru	46
5.3.3	Sodelovanje sledilnikov	51
5.3.4	Uporabniški vmesnik	52
5.3.5	Večnost sistema	52
6	Vrednotenje sistema in rezultati	55
6.1	Izbira parametrov	57
6.2	Scenariji eksperimentov	57
6.2.1	Scenarij 1: Gibanje dveh oseb pri konstantni osvetljenosti	59
6.2.2	Scenarij 2: Interakcija dveh oseb pri konstantni osvetljenosti in gibanje po zatemnjenem prizorišču	59
6.2.3	Scenarij 3: Izmenično gibanje in mirovanje dveh oseb z vmesnimi zatemnitvami	59
6.2.4	Scenarij 4: Pojavljanje in izginjanje dveh oseb z uvodno zatemni- tvijo in vmesnimi spremembami osvetljenosti	60
6.3	Rezultati	61
7	Zaključek	65
7.1	Glavni prispevki	65
7.2	Možnosti nadaljnjega raziskovanja	66

A Priloge	79
A.1 Slikovno zaporedje sledenja - 1. scenarij	80
A.2 Slikovno zaporedje sledenja - 2. scenarij	81
A.3 Slikovno zaporedje sledenja - 3. scenarij	82
A.4 Slikovno zaporedje sledenja - 4. scenarij	83

1 Uvod

1.1 Motivacija

Kdor se ukvarja s področjem sledenja objektov zagotovo ve, kakšen trud je vložen v načrtovanje metod, ki bi omogočale zanesljivo (dolgoročno) sledenje objektov in hkrati preprečevale odpoved sledenja v kriznih situacijah. Znano je, da lahko večina sledilnih algoritmov tekom sledenja hitro izgubi stik z realnostjo, predvsem zaradi neugodnih razmer in drugih problematičnih okoliščin, ki otežujejo ocenjevanje položaja sledenih objektov. Če se poglobimo v podrobnosti dotične implementacije, lahko hitro izvemo, kaj je razlog za odpoved določenega algoritma.

Namen naše raziskovalne naloge je izboljšati kvaliteto sledenja z uporabo dveh sledilnih algoritmov, ki operirata v ločenih domenah. Pri tem domeni zajemata področje infrardečega (IR) spektra in področje vidnega spektra. Povsem enostavno povedano – gre za dva opazovalca, ki na enako sceno gledata skozi različne oči. V teoriji to pomeni, da se sledilnik iz izbrane domene v primeru težav lahko obrne na “svojega sorodnika” v drugi domeni. Velja seveda tudi obratno. Lahko bi rekli, da želimo oba sledilnika (ki sta karakteristično zelo različna) pripraviti do tega, da se medsebojno dopolnjujeta in

koherentno usklajujeta proces sledenja.

Sistem omenjenih združenih sledilnikov bi bil praktično uporaben na različnih področjih. V našem primeru smo se osredotočili na sledenje in lokalizacijo oseb. Da bi sistem lahko preizkusili celostno in v najbolj neugodnih pogojih, je poudarek naloge tudi na sledenju v drastično spreminjajočih se svetlobnih pogojih. Zaradi tega vnaprej vemo, da bo sledilnik v vidnem spektru zagotovo imel velike težave pri prepoznavanju oseb s pomočjo barvne informacije. Kot oporo zato dodamo sledilnik v IR spektru, ki bo deloval v navezi s preprostimi točkastimi oznakami. Te oznake bodo pritrjene na osebe, ki jim želimo slediti. Za nameček pa bo naš sistem skušal slediti večim osebam ter tekom sledenja skušal ohranjati tudi njihovo identiteto.

1.2 Prispevek naloge

Prispevek naloge zajema poskus združitve dveh relativno preprostih sledilnikov, z namenom izboljšanja delovanja v spreminjajočih se svetlobnih pogojih. V ta namen implementiramo dva neodvisna sledilnika, od katerih eden operira v vidnem (barvnem) spektru, drugi pa v navezi z aktivnimi oznakami deluje v infrardečem (IR) spektru. Razlog za uporabo spektra IR je povsem preprost. V primeru, da globalna osvetlitev prostora postane šibka, ali kako drugače moteča, barvni sledilnik sicer odpove, vendar ga v tem času lahko nadomesti sledilnik IR, ki naj bi poskrbel za nemoteno delovanje sistema. Po drugi strani sledilnik IR ni popoln, saj zaradi točkastih oznak nima veliko dodatne vizualne informacije, kar lahko povzroči zamenjavo identitet vpletenih objektov. V tem primeru pričakujemo, da se bo identiteta tarč ohranjala glede na vizualno informacijo iz barvnega sledilnika. V nalogi izpostavimo tudi težave, ki so nastale pri implementaciji sistema ter podamo iztočnice za nadaljnje delo na tem področju.

1.3 Metodologija

Ker je področje naloge izredno obširno, metode v poglavju kategorizacije najprej razdelimo glede na splošen pristop reševanja problematike. Tako dobimo dve glavni kategoriji – verjetnostne metode in metode determinističnega tipa (optimizacije). V nadaljevanju pregleda področja zatem predstavimo tudi konkretne pristope, ki so del algoritmov za sledenje – tako zajamemo številne modele za modeliranje izgleda, gibanja, prekrivanja, interakcije ter izključevanja. Pregled zaključimo z omembo prepoznavnih metod,

ki združujejo nekatere predhodno opisane modele. V praktičnem delu zatem natančneje opišemo uporabljene algoritme in postopke, ki sestavljajo osnovno implementacijo ter predstavimo tudi preostali programski del (arhitekturo), ki združuje celoten sistem.

1.4 Pregled naloge

V poglavju 2 problem sledenja objektov matematično formuliramo. Formulacija služi kot okvir za nadaljevanje pregleda na področju sledilnih algoritmov. Zatem v poglavju 3 metode kategoriziramo glede na njihove poglavitne lastnosti in področje uporabe. V poglavju 4 najdemo glavni del pregleda obstoječega področja. Poglavje je strukturirano na komponente, iz katerih so sestavljeni tipični sledilni algoritmi. V kolikor je to potrebno, so določene razlage podprte z matematičnimi opisi iz izvornih del. V poglavju 5 najdemo opis implementacije našega sledilnega sistema, poleg tega pa tudi študijo primerov uporabe ter popis težav, ki so se pojavile pri implementaciji. V poglavju 6 prikažemo vrednotenje implementiranega sistema na štirih preizkusnih scenarijih. Zaključno poglavje 7 ponudi komentar na opravljeno delo, izpostavi nerešene težave in predlaga možne izboljšave.

1.5 Terminologija

V nadaljevanju opišemo ključne izraze s katerimi se srečujemo na področju sledenja večih objektov z namenom, da bi bil preostanek naloge bralcu bolj razumljiv.

Objekt: V računalniškem vidu je objekt na sliki predstavljen kot zaprto področje slikovnih elementov, ki se prepoznavno razlikuje od ozadja. V primeru sledenja večih objektov so objekti lahko vizualno zelo podobni oz. enakega tipa, zato se pojavi potreba po individualnem označevanju objektov oz. njihovi identifikaciji.

Detekcija: Detekcija predstavlja lokalizacijo predhodno opisanih objektov v slikah. Naloga detektorja torej je, da nam pove, kje na sliki se iskani objekt nahaja (lahko tudi njegovo velikost ali kako drugo lastnost). Običajno je pojem detekcije zelo povezan s področjem strojnega učenja, saj se mora detektor pogosto najprej naučiti (iz obstoječe učne množice), kako objekt izgleda. S tem detektor naučimo razlikovati med entitetami, ki predstavljajo naš objekt in drugimi objekti oziroma ozadjem. Detekcija objektov običajno ne vsebuje časovne informacije – ta nastopi šele pri sledenju objektov.

Sledenje: Sledenje je lokalizacija objekta v zaporedju slik – zaradi tega vsebuje

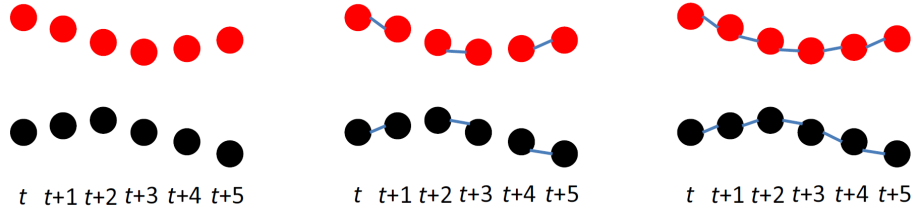
časovno informacijo. Pri sledenju večih objektov to pomeni lokalizacijo vseh sledenih objektov ter ohranjanje njihove identitete tekom sledenja.

Odziv detektorja: Odziv detektorja (hipoteze detektorja) so rezultati, ki jih vrne detektor, ki je bil naučen prepoznavati specifične objekte (ljudi, vozila, obraze, živali). Rezultati lahko vsebujejo pozicije objektov, dimenzije in po možnosti tudi druge opisne lastnosti pridobljene iz analizirane slike.

Trajektorija: Trajektorije so končni rezultat sledilnega sistema. Pri tem vsaka posamezna trajektorija pripada svojemu objektu. Sama trajektorija je sestavljena iz množice detekcij objekta (tarče) v zaporedju slik, kjer vsaka detekcija predstavlja lokacijo, velikost in druge lastnosti tarče na eni izmed slik zaporedja.

Sled (angl. *tracklet*): Sled je vmesni rezultat sledenja. Sestavljena je iz zaporednih napovedi, za katere se predvideva, da pripadajo istemu objektu. Sled s tem postane del trajektorije tega objekta oz. sama po sebi predstavlja delno trajektorijo. V nekaterih pristopih se trajektorije kreirajo s postopnim lepljenjem detekcij v sledi, dokler se s temi sledmi eventuelno ne sestavi celotna trajektorija gibanja objekta.

Združevanje podatkov (angl. *data association*): Združevanje detekcij med seboj je tipičen pristop pri sledenju objektov. Pri tem je naloga združevanja najti vse korespondence oz. ujemanja med detekcijami objektov v zaporednih slikah – tako, da lahko na ta način sestavimo sledi in kasneje tudi trajektorije.



Slika 1.1 Odzivi detektorja (levo), delne sledi (v sredini) in trajektorije (desno) v 6 časovnih korakih. Barvno kodiranje predstavlja dva ločena objekta [36].

2 Formuliranje problema

Rezultat sledenja objektov so trajektorije vseh objektov, ki se v zaporedju slik premikajo po slikovni ravnini. V nadaljevanju predstavimo matematično formulacijo tega problema.

Sledenje večim objektom lahko definiramo kot problem ocenjevanja stanja večih spremenljivk. Kot vhod lahko vzamemo zaporedje slik $\{I_1, I_2, I_3, \dots, I_t, \dots\}$ in definiramo stanje s_t^i , ki nam anotira stanje i -tega objekta v sliki s časovno oznako t . Zatem lahko definiramo množico vseh stanj v času t kot $S_t = \{s_t^1, s_t^2, s_t^3, \dots, s_t^{M_t}\}$ za vseh M_t objektov v pripadajoči sliki I_t . Množico vseh stanj tekom celotnega zaporedja pa lahko zapišemo kot $S_{1:t} = \{S_1, S_2, \dots, S_t\}$. Pri tem upoštevamo, da se lahko število objektov M_t spreminja skozi čas.

Da bi lahko ocenili stanja objektov v času t , je potrebno iz slik pridobiti opazovanja v obliki meritev oz. detekcij. Definiramo lahko o_t^i , ki predstavlja opazovano meritev i -tega objekta v času t . Zatem lahko za sliko v času t definiramo množico opazovanih meritev kot $O_t = \{o_t^1, o_t^2, o_t^3, \dots, o_t^{M_t}\}$. Prav tako lahko množico vseh opazovanih meritev opišemo kot $O_{1:t} = \{O_1, O_2, \dots, O_t\}$ (od prve slike do slike z indeksom t , za vse objekte).

Sledilni algoritem mora zatem z uporabo vhodnih meritev najti optimalna zaporedna

stanja vseh objektov. Iskanje optimalnosti lahko modeliramo kot ocenjevanje maksimalne aposteriorne verjetnosti (MAP) iz pogojne porazdelitve zaporednih stanj vseh objektov, če imamo na voljo vsa opazovanja [45]. To pomeni, da preiščemo celoten prostor (porazdelitev) stanj in opazovanj in najdemo najbolj verjetna stanja, glede na meritve iz vseh opazovanj. To lahko zapišemo v obliki:

$$\hat{S}_{1:t} = \arg \max_{S_{1:t}} P(S_{1:t} | O_{1:t})$$

Optimizacijo lahko izvedemo z rekurzivnimi verjetnostnimi pristopi, ki bazirajo na dveh korakih [36]. V prvem koraku se izvede predikcija oz. napoved stanja glede na prejšnja opazovanja (brez zadnjega), kar lahko zapišemo kot:

$$P(S_t | O_{1:t-1}) = \int P(S_t | S_{t-1}) P(S_{t-1} | O_{1:t-1}) dS_{t-1}$$

V drugem koraku se izvedejo popravki stanj, glede na vsa prejšnja opazovanja:

$$P(S_t | O_{1:t}) \propto P(O_t | S_t) P(S_t | O_{1:t-1})$$

Pri tem $P(S_t | S_{t-1})$ predstavlja dinamični model (angl. *dynamic model*), $P(O_t | S_t)$ pa model opazovanja (angl. *observation model*). Verjetnostni porazdelitvi teh modelov sta običajno neznani, zato se za oceno teh porazdelitev pogosto uporabijo različne metode vzorčenja. Kot bomo videli v nadaljevanju, ta dva modela kritično vplivata na izvedbo algoritmov sledenja.

Drugi način iskanja optimalnih stanj objektov so deterministične metode optimizacije. To pomeni, da poskusimo direktno najti maksimum verjetnostne funkcije $P(O_{1:t} | S_{1:t})$ z gradientnim vzponom in s tem pridobiti optimalno rešitev. Ekvivalentno lahko najdemo maksimum verjetnostne funkcije $P(S_{1:t} | O_{1:t})$. Optimizacijo v tem primeru lahko zapišemo kot:

$$\hat{S}_{1:t} = \arg \max_{S_{1:t}} P(S_{1:t} | O_{1:t}) = \arg \max_{S_{1:t}} P(O_{1:t} | S_{1:t})$$

Konkretnih možnih pristopov je več, npr. iskanje ujemanja dvodelnih grafov, dinamično programiranje, maksimalni pretok z minimalno ceno, neodvisna množica maksimalnih uteži [36]. Več o teh metodah sledi v zadnjem delu poglavja 4.

3 Kategorizacija metod

V splošnem je metode sledenja objektov zaradi velike raznolikosti težko enolično kategorizirati. V obstoječih delih lahko naletimo na določene poglede glede klasifikacije metod (npr. glede na uporabo značilk, predstavitve objektov ipd.) [36, 71]. Zaradi potrebe po sistematičnemu pristopu se zgledujemo po kategorizaciji iz [36], ki metode ločuje glede na več ključnih razlik, ki jih opišemo v nadaljevanju.

3.1 Glede na inicializacijo

Metode sledenja lahko ločimo glede na način inicializacije objektov [67]. Pri tem obstajata dve kategoriji metod. Prva je sledenje, ki bazira na detekciji objektov (DBT), druga pa sledenje, ki te detekcije ne potrebuje (DFT). V nadaljevanju opišemo obe kategoriji.

3.1.1 Sledenje z detekcijo objektov

Pri tem načinu sledenja objekte detektiramo (lokaliziramo) v vsaki sliki (primer v [56]). Zatem pozicije objektov iz detekcij povežemo v trajektorije gibanja. Objekte lahko detektiramo z rekonstrukcijo gibanja v zaporednih slikah (modeliranje ozadja) ali pa s

posebnim detektorjem, ki prepoznava specifično vrsto iskanih objektov. Pri tem je sledenje lahko sprotno (angl. *sequential*) ali paketno (angl. *batch processing*) [36]. V večini primerov se tukaj uporabijo vnaprej pripravljeni detektorji, npr. diskriminativni detektor iz [12]. Tovrstni detektorji v splošnem dobro prepoznava določeno vrsto objektov (npr. osebe, vozila, obraze in podobno). Težava vnaprej pripravljenih detektorjev je predvsem ta, da njihova zmogljivost neposredno vpliva tudi na kvaliteto hipotez in trajektorij sledenja. Konkretno, če zelo splošen detektor uporabimo na zelo specifičnem problemu, obstaja velika verjetnost, da bo deloval slabo in tako negativno vplival na pravilnost sledenja.

3.1.2 Sledenje brez predhodne detekcije

Če sledimo objekte brez predhodne detekcije, je še vedno potrebno definirati, kje se omenjeni objekti v sliki nahajajo. Običajno se to naredi ročno na začetku algoritma, kar pomeni, da objekte označimo na prvi sliki, ki jo dobi algoritem na vhodu (primer take inicializacije najdemo v [21]). Prednost je predvsem ta, da s tem lahko sledimo poljubne objekte. Algoritem se v nadaljnjem procesiranju ne zanaša na detektor. Glavna omejitev sicer je, da se število objektov tekom sledenja ne spreminja, zato izginjanje in pojavljanje novih objektov predstavlja veliko težavo, saj teh problemov ta pristop ne rešuje. S tem se zmanjšajo tudi možne praktične aplikacije tega pristopa, sam pristop pa postane zelo podoben klasičnemu sledenju predlog (angl. *template tracking*).

3.2 Glede na način procesiranja

Metode lahko ločimo glede na to, kako se vhodni podatki procesirajo. Kot je bilo že omenjeno, ločimo sprotni in paketni pristop. Ta kategorizacija je predstavljena v [36]. Sorodna izraza, ki ju najdemo v literaturi sta tudi sledenje v realnem času (angl. *online tracking*) in sledenje, ki se ne dogaja v realnem času (angl. *offline tracking*). Glavna razlika je v tem, kako obravnavamo vhodne podatke. Pri sledenju v realnem času nimamo na voljo vseh sličic, zato je to sprotno sledenje. Pri paketnemu sledenju je zaporedje slik vnaprej znano, kar pomeni, da v vsakem koraku sledenja lahko pogledamo tudi v prihodnost. To seveda pomeni, da se sledenje v tem načinu ne izvaja v realnem času, saj analiziramo obstoječ video posnetek. V teoriji bi zato paketni pristopi lahko našli globalno optimalno rešitev, ker lahko naenkrat pregledajo vse vhodne slike, kar pa seveda ne drži za sprotne metode.

3.2.1 Sprotno procesiranje

Kot rečeno imajo metode sprotnega tipa vpogled le v minule slike v zaporedju, ne morejo pa pregledovati prihajajočih slik. Sledenje se zato izvaja koračno (za vsako sliko posebej), kar pomeni, da se nove pozicije objektov generirajo sproti in se dodajajo k obstoječim trajektorijam. Primerne so za realnočasovne aplikacije, glavna pomanjkljivost pa pri napovedovanju povzročajo nepopolne meritve oz. manjkajoče detekcije iz preteklosti. V praksi so metode sprotnega sledenja večinoma verjetnostnega tipa, čeprav obstajajo tudi deterministični pristopi. Primeri v [6, 7, 24, 60]

3.2.2 Paketno procesiranje

Paketno procesiranje zajame meritve iz vseh slik. Meritve so analizirane skupaj z namenom, da bi se sestavile optimalne globalne trajektorije potovanja objektov, zaradi česar je pristop zelo popularen (najdemo ga velikokrat – [8, 20, 30, 48, 56, 68]). V primeru, da gre za računsko zahtevno analizo, se video zaporedje lahko razdeli na več segmentov. Procesiranje teh segmentov ostaja še vedno paketno, na koncu pa je potrebno delne rešitve hierarhično združiti v globalno rešitev. Slabost pristopa je v zakasnjem pridobivanju rezultatov, kar pomeni, da ni primeren za realnočasovne aplikacije. Zaradi poskusa iskanja globalno optimalnih parametrov se običajno uporabljajo deterministične metode.

3.3 Glede na matematično ozadje

Pri formulaciji problema sledenja smo že izpostavili dva tipa optimizacijskih pristopov – verjetnostnega in determinističnega. Luo et al. [36] metode kategorizira tudi na ta način.

Tukaj omenimo dve glavni razliki, ki se v teh pristopih tipično pojavljajo. Prva razlika je bila na kratko že omenjena – to je razlika v samem pristopu reševanja problema. Vsak pristop namreč ocenjuje stanja objektov na svoj način. Pri verjetnostnih metodah se ocena pridobi z verjetnostnim sklepanjem, medtem ko se pri determinističnih metodah uporabijo pristopi za optimizacijo.

Druga razlika pa je oblika, v kateri dobimo končne rezultate. Pri tem nam lahko verjetnostni pristop zaradi vključenega faktorja naključnosti pri vzorčenju verjetnostne porazdelitve generira nekoliko različne rešitve pri vsakem izvajanju nad istimi podatki. Deterministični pristop bo vedno generiral enake rešitve, ker končna rešitev ne izhaja iz

verjetnostne porazdelitve.

3.3.1 Verjetnostne metode

Kot je bilo že omenjeno se pri verjetnostnih metodah stanje objektov pridobi z iterativnim ponavljanjem dveh ključnih korakov (napovedi in posodobitve). V praksi se pogosto zgodi, da stanja objektov ne moremo vedno zadovoljivo opisati z neko splošno uporabno porazdelitvijo (npr. pogosto uporabljeno normalno porazdelitvijo). Zaradi tega se lahko pojavijo težave v koraku napovedovanja, kjer analitična rešitev ni vedno mogoča. Dodaten problem predstavlja še večje število sledenih objektov, kar posledično zviša tudi dimenzije stanj, ki jih ocenjujemo in otežuje izračun koraka napovedovanja. Primeri pristopov so navedeni v poglavju 4.6, kjer pogosto uporabljene metode tudi natančneje opišemo.

3.3.2 Deterministične metode

V praksi deterministične metode ponavadi delujejo bolje od verjetnostnih metod, še posebej v primerih izrazitega prekrivanja opazovanih objektov. Z možnostjo uporabe globalne informacije (celotne video sekvence) se namreč prekrivanje lahko rešuje bolje kot pri nekaterih verjetnostnih metodah. Po drugi strani dostopnost globalne informacije zahteva več časa in prostora za izvedbo optimizacije ter vnaprej znano sekvenco slik, zaradi česar so deterministične metode velikokrat omejene na bločno procesiranje. Primeri in opisi pogostih pristopov so navedeni v poglavju 4.7.

3.4 Glede na področje uporabe

Kategorizacija metod je možna tudi glede na področje njihove uporabe [36]. Tukaj navedemo nekatera najbolj pogosta področja uporabe sistemov za sledenje objektom. Velikokrat je namreč težko enolično kategorizirati pristope le glede na uporabljeno metodologijo, zato ta razdelitev predstavlja alternativen način kategorizacije.

3.4.1 Športni scenariji

Sledenje na igrišču je uporabljeno za pridobivanje statističnih podatkov na tekmah ali pa avtomatsko analizo igralnih strategij. Običajno imamo različno postavitev večih kamer, kjer so posnetki včasih tudi zelo dinamični, kar problem sledenja oteži. Po drugi strani pa je vizualna meja med igralno površino in ostalimi regijami v določenih športih

(nogomet, košarka, tenis) dokaj jasno določena. To pomeni, da lahko učinkovito modeliramo ozadje, tako kot se tega v svojem delu poslužuje Xing et al. [64], kjer je ločevanje površine igrišča od ostalih površin ključen korak predprocesiranja. Druga značilnost tega področja je, da so objekti, ki jih sledimo lahko obenem barvno zelo podobni, hkrati pa se močno razlikujejo od ozadja – še posebej, če sledimo npr. igralcem med tekmo. Zaradi barvne podobnosti je težje ohranjati identiteto posameznih objektov, lahko pa olajša npr. ločevanje igralcev glede na ekipo. Primer najdemo v delu [35], kjer Lu et al. predlaga klasifikator s pomočjo logistične regresije, ki preslika slikovne zaplate v labela ekip z uporabo barvnih histogramov. Še en primer analize športne tekme je na voljo v [27], kjer Kristan et al. uporabi več sledilnikov, ki so namenjeni sledenju ene tarče (angl. *single target trackers*), za ohranjanje identitet pa predlaga metodo s parceliranjem regij *Voronoi* (angl. *Voronoi partitioning*).

3.4.2 Letalstvo in scenariji iz zraka

V to kategorijo spadajo vsi pristopi, ki analizirajo zajete posnetke iz zraka. Običajno se v teh posnetkih giblje večje število tarč, ki se nahajajo na tleh (npr. analiza in nadzor vozil). Pri tem sledenju so lahko obravnavani objekti zelo majhni, kar pomeni, da uporaba klasičnih vizualnih modelov za njihovo opisovanje ni primerna. Hkrati je zgoščenost tarč na sliki lahko velika, kar pomeni, da gre za sledenje večjega števila objektov. Primer najdemo v [49], kjer se za detekcijo gibanja uporablja modeliranje ozadja. V vsakem koraku sledenja se zatem za dodeljevanje objektov k trajektorijam uporabi t.i. Madžarska metoda (angl. *Hungarian algorithm*), ki poišče najbližje globalne sosedje. Pri za ohranjanje identitete se upoštevajo parametri, ki jih lahko pridobimo z opazovanjem gibanja objektov (npr. afiniteta, atrakcija, hitrost). Še primer analize iz zraka najdemo v [53].

3.4.3 Sledenje pešcev in prometa

Področje detekcije in sledenja pešcev in vozil se večkrat uporablja za namene nadzora, veliko zanimanje pa je tudi v avtomobilski industriji, kjer v razvoju poteka avtomatizacija različnih nalog vozila. Ocenjuje se, da je velika večina oz. kar 70 odstotkov vsega raziskovalnega dela na področju detekcije namenjenega prav detekciji pešcev in vozil. Kot je bilo že omenjeno, so to največkrat vnaprej naučeni detektorji, ki lahko, glede na svoje predispozicije, zelo vplivajo na zmogljivost sledenja. Primere sledenja pešcev v prometu najdemo v [2, 13, 17, 25, 33, 38, 42].

3.4.4 Splošna uporaba

Tukaj lahko omenimo določene hibridne pristope, ki so se izkazali kot dokaj uspešni. Zhao et al. [74] predlaga sledenje, kjer je labeliranje objektov zahtevano v prvi sliki, šele zatem se tem objektom začne slediti, hkrati pa se v prvih nekaj slikah detektor nauči, kako označeni objekti izgledajo. S tem se pridobi ustrezno podporo detektorja za prepoznavanje označenih objektov v nadaljevanju. Čeprav ima tudi ta pristop omejitvev v smislu konstantnega števila objektov, še vedno predstavlja zanimivo združevanje ročnega označevanja in avtomatske detekcije, hkrati pa zmanjša količino ročnega dela, ki ga je potrebno opraviti pri izključno ročnem označevanju.

Bolj praktično generalizacijo v svojem delu predstavita Luo in Kim [62], kjer sledita več podobnih objektov na podlagi le enega označenega primera v prvi sliki. Detekcija objektov v naslednjih slikih se postopoma izoblikuje preko akumulacije vseh učnih primerov.

4 Komponente sledenja

V splošnem vsako metodo sledenja sestavljata dve glavni komponenti – model opazovanja in dinamični model [36]. Model opazovanja služi kot orodje za primerjavo oz. oceno podobnosti med stanji in novimi opazovanji oz. meritvami. Konkretno to pomeni, da mora model opazovanja modelirati vizualni izgled objekta, njegovo gibanje, medsebojno interakcijo, izključevanje in prekrivanje. Dinamični model pa modelira prehajanje stanj med posameznimi koraki sledenja oz. povezuje stanja objektov iz slike v sliko. Kot smo že omenili, tudi tukaj lahko modele delimo na verjetnostne in deterministične. V nadaljevanju opišemo koncepte, ki jih naslavljata oz. rešujeta obe komponenti.

4.1 Vizualni modeli

Ko govorimo o vizualnih opisih objektov, sta v ospredju dve komponenti – vizualna predstavitev objekta in zajem statističnih lastnosti te predstavitve. Pri tem je vizualna predstavitev močno povezana s pojmom značilk, vendar običajno poleg izbire značilk zajema tudi druge parametre. Glavni pomen vizualne predstavitve je čim bolj natančen opis vizualnih karakteristik. Možna je predstavitev le z enim tipom značilk (enostavni

modeli), ali pa gre za uporabo in združevanje večih tipov značilk (sestavljene modeli).

Statistične lastnosti pa (kot druga komponenta) označujejo orodja, ki omogočajo primerjavo (mero podobnosti) med objekti z danimi vizualnimi predstavitvami. Formalno lahko to zapišemo kot $S_{ij} = F(o_i, o_j)$, kjer je F funkcija, ki vrača podobnost, o_i in o_j sta vhodna parametra, ki predstavljata statistično meritev opazovanj i in j , S_{ij} pa predstavlja mero ujemanja teh opazovanj [36]. V nadaljevanju opišemo pogosto uporabljene vizualne značilke ter njihovo uporabo v enostavnih ter sestavljenih vizualnih modelih.

4.1.1 Vrste značilk

Brez ustreznih značilk pri sledenju objektov ne gre, saj z njimi opišemo, prepoznavamo in primerjamo objekte. Tukaj vrste značilk ustrezno kategoriziramo in opišemo.

Značilne točke

Konkreten primer uporabe značilnih točk predstavi Sugimura et al. [57], kjer so uporabljene značilke KLT [61], s pomočjo katerih se generira sledi oz. trajektorije. Podobno so za zajemanje teksturnih karakteristik uporabljene točke v obliki lokalnih značilnic [69]. Ena izmed različic točkastih značilk pa je uporabljena tudi za gručenje vzorcev gibanja [9].

Barvne in intenzitetne značilke

Značilke v tej kategoriji so v sledenju najpogostejše uporabljene. Lahko so predstavljene v surovi obliki slikovnih elementov (slikovna predloga) [66] ali pa v obliki histogramov [23, 40, 56, 57]. Pri tem lahko vsebujejo informacijo o intenziteti ali pa konkretnih barvnih vrednostih (odvisno od tega ali analiziramo sivinsko ali barvno sliko). Glavna funkcionalnost je primerjava podobnosti med meritvami. Histogram pri tem ne ohranja prostorskih značilnosti, ker gre za statistično predelano obliko (slikovne elemente le preštejemo in jih združujemo glede na njihovo vrednost). Slikovna predloga (izrezana regija) po drugi strani ohranja prostorske značilnosti, vendar je zelo rigidna in deluje dobro le pri objektih, ki se vizualno ne spreminjajo.

Optični tok

Optični tok se uporablja predvsem za kratkoročno sledenje. V določenih rešitvah je uporabljen za povezovanje posameznih detekcij v krajše sledi, kar lahko predstavlja predho-

dni korak pred dejanskimi asociacijskimi metodami, ki sledi združijo v trajektorije [50]. Spet drugje se neposredno uporablja v podatkovni asociaciji [23]. Poleg tega v nekaterih primerih lahko dopolnjuje druge obstoječe značilke, ki tvorijo opazovalni model [3]. Uporablja pa se tudi v scenarijih z veliko gostoto objektov, kjer se gibanje gneče modelira s socialnimi modeli gibanja [1].

Gradientne značilke

Sem spadajo histogrami usmerjenih gradientov (HOG), ki so najbolj prepoznavno uporabljeni pri detekciji ljudi in sledenju pešcev [12]. Uporaba značilk HOG je zelo številčna [7, 23, 30]. V to kategorijo spadajo tudi vsi podobni pristopi, ki značilke pridobivajo s primerjavo sosednih slikovnih elementov. Mitzel et al. [40] uporablja detektor HOG, obenem pa z uporabo metode nivojskih množic (angl. *level-set method*) izvaja tudi segmentacijo dobljenih detekcij.

Regijske kovariančne matrice

Porikli et al. predstavi regijske kovariančne matrice v [44]. Njihova značilnost je odpornost tako na spremembe v osvetlitvi, kot tudi na spremembe v skaliranju objektov. Uporabljene so pri identifikaciji in vizualni asociaciji tarč [20], lahko pa predstavljajo osnovo za učenje vizualnih informacij oz. način predstavitve vizualnega dela objektov [21, 30].

Globinska informacija

Uporabljena je na različnih področjih računalniškega vida. Mitzel et al. [40] globinsko informacijo uporablja kot orodje za popravljanje in ponovno inicializacijo pravokotnika, ki omejuje regije objektov (angl. *bounding box*). Ess et al. [13] v svojem delu izboljšuje detektiranje objektov na mobilnem sistemu, ki izvaja sledenje objektov. Globina prav tako pomaga pri sledenju tarč v scenariju s premikajočim se vozilom, kot ga predstavita Gavrila in Munder [17]. Služi pa lahko tudi kot podpora 3D sledenju v [18], kjer se na ta način ocenjuje uteži delcev pri sledenju s filtrom z delci.

Ostalo

Sem spadajo npr. značilke na podlagi hoje (angl. *gait features*) iz frekvenčne domene, ki so unikatne za vsako osebo [57]. S njihovo pomočjo Sugimura et al. maksimizira

zmožnost razločevanja sledenih oseb. Pri tem se iz vsake trajektorije sistem z uporabo linearne regresije nauči, kako izgleda njena periodična komponenta. Nad to komponento se zatem izvede hitra Fourierova transformacija (FFT), s katero se pridobi amplituda in faza trajektorije. Vsako trajektorijo se tako lahko loči od ostalih na podlagi podobnosti, ki je izračunana iz teh dveh lastnosti.

Zemljevid verjetnostne zasedenosti (POM), ki ga najdemo v [4] in v [16], je namenjen ocenjevanju verjetnosti, da se bo objekt nahajal na specifičnih koordinatah vnaprej pripravljene mreže pri sledenju z uporabo večih kamer. Na podlagi modeliranja ozadja postavi več hipotez o lokaciji objektov, kar predstavlja vhodne detekcije v sistem sledenja večih oseb.

V splošnem lahko rečemo, da je večina značilk dokaj učinkovitih. Hkrati velja omeniti, da ima večina tudi omejitve in različne slabosti, zato niso vse vrste značilk primerne za reševanje poljubnih problemov. Kot je bilo že omenjeno, je npr. barvni histogram že preizkušen in dobro poznan način merjenja podobnosti dveh regij, vendar se moramo zavedati, da pri tem ne ohranja prostorske informacije. Do težav lahko prihaja v primeru, da imamo dva barvno podobna objekta, ki se po obliki oz. prostorski predstavitvi občutno razlikujeta. Značilne točke so lahko učinkovite, vendar slabo rešujejo prekrivanje ali prostorske rotacije izven slikovne ravnine. Gradientne značilke so izredno robustne na spremembe v svetlobi oz. osvetljenosti, ker opisujejo prehode oz. robove, vendar so prav tako lahko občutljive na prekrivanje ter tudi na deformacije objekta. Regijske kovariančne matrike so še robustnejše, ker skušajo zajeti še več informacij, vendar so zato tudi računsko zahtevnejše. Uporaba globinske informacije lahko izboljša natančnost, vendar zahteva več pogledov na isto sceno ter tudi dodatne algoritme za kalibracijo, procesiranje in zajemanje globinskih vrednosti.

4.1.2 Enostavni modeli

V večini primerov so enostavni modeli zaradi svoje učinkovitosti največkrat izbrani za sledenje objektov. V nadaljevanju so opisane različne izvedbe teh modelov.

Slikovna predloga

Predloga je najbolj enostaven model, saj gre za izrezano regijo slike, kjer je vsa informacija (intenziteta ali barvna vrednost) zapisana v izvornih slikovnih elementih. Prav zaradi tega slikovna predloga ohrani celotno prostorsko informacijo objekta. Primerjava

med dvema predlogama poteka po soležnih elementih, najbolj enostavna metoda za primerjanje je normalizirana križna korelacija (NCC) [66], kjer z drsečim oknom poiščemo najboljše možno ujemanje izvirne predloge (modela) in predloge na predvideni novi poziciji objekta v sliki. Pri tem se izvirna predloga (model) lahko med sledenjem postopoma osvežuje in prilagaja trenutnemu izgledu objekta [1].

Primer sledenja večih objektov v povezavi z NCC najdemo v [63]. Problem večih tarč je tukaj rešen s simuliranjem omrežnega pretoka. Križna korelacija je uporabljena za povezovanje detekcij po izračunu cen omrežnih prehodov. Podobno se v [43] avtorji poslužujejo enostavnega sledilnika s predlogami, ki je uporabljen za izračun verjetnosti nahajanja objektov glede na vizualni izgled. Podobnost je izračunana med začetno predlogo in predvideno predlogo kot eksponenten izraz $P_d(p) = \exp(-(NCC(p, p_0) - 1)^2)$, kjer je p predvidena predloga, p_0 začetna predloga, NCC pa operacija normalizirane križne korelacije. Čeprav je pristop učinkovit, mu težave predstavljajo tako sprememba osvetlitve kot tudi prekrivanje.

Histogram

Histogram je najbolj popularen vizualni model v sledenju, saj učinkovito hrani statistiko vizualnih značilnosti izbrane regije. Barvni histogrami, predstavljeni s strani Pereza et al. [47], so na primer uporabljeni za izračun funkcije podobnosti, pri tem pa je uporabljena eksponentna funkcija, ki podobnost histogramov pretvori v verjetnostno funkcijo [26]. Podobno je za določanje razdalje dveh histogramov lahko uporabljena razdalja *Bhattacharyya* (npr. pri konstruiranju grafa prehodov v [57]). Nastopa lahko v kombinaciji z Mean Shift algoritmom, kot je to realizirano v [11], kjer se s histogrami opišejo sekvenčne detekcije objektov. Histogram lahko nosi različno kodirano informacijo. Uporabo barvnega histograma s tremi barvnimi kanali (RGB) najdemo v [73], kjer so tako opisane posamezne detekcije. Na podoben način je definiran model v [33], kjer avtorji z njim definirajo vizualni model trajektorije. S prvo detekcijo objekta se histogram inicializira, zatem pa se z vsako naslednjo detekcijo razvije v uteženo povprečje histogramov trajektorije objekta. Pri tem se za mero podobnosti dveh histogramov uporabi že omenjena razdalja *Bhattacharyya*. Na podoben način je histogram z zapisom v barvnem prostoru HSV uporabljen v [48]. Avtorji v [65] s histogrami predstavijo posamezne prekrivajoče se regije oseb – osebo tako razdelijo na področje, ki zajema celo telo, na območje z glavo in rameni ter območje z glavo in trupom. Povezovanje trajektorij zatem izvedejo ločeno,

z združevanjem posameznih regij. Kljub različnim pristopom omejitev še vedno ostaja, saj se prostorska informacija tarče (oblika, robovi) v histogramu v nobenem primeru ne ohrani.

Kovariančna matrika

Metoda dobro deluje pri svetlobnih spremembah in spremembah rotacije objekta. Pristop je bil uporabljen v [20]. Modeliranje podobnosti je definirano kot

$$P_{link}(T_i, T_j) = \mathcal{N}(d_{ij}^a, \Sigma_a) \mathcal{N}(d_{ij}^m, \delta_t \Sigma_m) \mathcal{N}(d_{ji}^m, \delta_t \Sigma_m),$$

kjer sta T_i in T_j primerjani delni trajektoriji, \mathcal{N} predstavlja normalno porazdelitev, d_{ij}^a je mera vizualne različnosti (angl. *appearance dissimilarity*), d_{ij}^m je razlika med predvideno pozicijo trajektorije T_i in najbližjo detekcijo, ki pripada trajektoriji T_j , δ_t je časovni korak, parametra Σ_a in Σ_m pa sta pripadajoči kovariančni matriki, ki se ocenita iz učne množice. Obstaja tudi izboljšava, ki regijo objekta razdeli na več blokov [21]. Tukaj se kovariančna matrika generira za vsak blok posebej. Iz soležnih blokov izvirne regije in detektirane regije se določi funkcija podobnosti. Mera podobnosti celega objekta je produkt podobnosti vseh blokov.

Binarna primerjava slikovnih elementov

Tukaj je mišljeno neposredno primerjanje dveh soležnih slikovnih elementov, kjer na izhodu dobimo binaren rezultat. Zhao et al. [74] predstavi primer uporabe naključnih gozdov (angl. *random ferns*), ki iz učne množice izračunajo rezultate primerjav in s pomočjo glasovanja rezultate tudi klasificirajo. Pri tem glede na učne podatke vsako novo regijo klasificirajo v pozitivno ali negativno skupino (glede na delež pozitivnih in negativnih vzorcev v učnih primerih).

Vreča pojmov

Vreča pojmov (angl. *bag of words*) predstavlja način hranjenja vnaprej izračunanih značilk. Tako so na primer lokalne značilke (SIFT) zapisane na ta način v [69]. Ker zapis značilk sam po sebi ni dovolj za vizualno predstavitev objekta, se ohrani tudi prostorska informacija z uporabo metode prostorskega piramidnega ujemanja (SPM) [32].

4.1.3 Sestavljeni modeli

Ob prepričanju, da združevanje modelov izboljša robustnost in kompenzira pomanjkljivosti posameznih metod so se pojavili tudi sestavljeni modeli. Problem teh modelov je večja kompleksnost. Pojavi pa se tudi vprašanje učinkovitega združevanja informacij iz večih enostavnih modelov. Tukaj je predstavljenih nekaj pristopov združevanja, ki se uporabljajo pri realizaciji sestavljenih modelov. Kategorizacija je povzeta iz [36].

Boosting

Boosting (krepitev) je strategija izbiranja tistih značilk, ki imajo najboljše razločevalne (diskriminativne) lastnosti. Obstaja več algoritmov, najbolj znana sta *AdaBoost* [30] in *RealBoost* [67]. Diskriminativni vizualni model, ki povzema to strategijo je bil predstavljen v [30]. Izbrane značilke modela sestavljajo barvni histogram, gradientne značilke (HOG) in deskriptor iz kovariančnih matrik. Z algoritmom *AdaBoost* avtorji skušajo izbrati tiste značilke, ki najboljše opišejo ujemanje podobnih objektov in hkrati ločujejo trajektorije, ki pripadajo različnim objektom. Enake značilke so uporabljene v [67], kjer so kombinirane z algoritmom *RealBoost*.

Predlagan je bil tudi algoritem *HybridBoost* [34] z namenom avtomatične izbire najbolj diskriminatornih značilk. Pri tem si pomaga s t.i. funkcijo izgube, ki je sestavljena iz dveh delov. Klasifikacijski del kaznuje napačne asociacije, del za rangiranje pa daje višjo vrednost pravih asociacijam.

Konkatenacija

Ta primer združevanja značilnic s konkatenacijo najdemo v [8]. Značilke so podlaga za klasifikator SVM, ki ločuje izbran objekt od ostalih objektov v določenem časovnem okviru. Zajemajo barvno in gradientno informacijo (HOG) ter optični tok. Konkatenacija se izvede z redukcijo dimenzionalnosti, konkretno z metodo analize glavnih komponent (PCA).

Seštevanje

Ta način združevanja značilk je uporabljen v [40], kjer so obravnavani primeri zelo podobnih barvnih odtenkov pešcev in ozadja. Za reševanje tega problema se zato skupaj z barvnimi značilnicami uporabi tudi globinska informacija. Za objekte v ospredju se izračuna pričakovana globina glede na njihovo največjo hitrost in trenutno globino. V

naslednjih korakih lahko iz pričakovane globine določimo verjetnost naslednje globine vsakega objekta. Ta verjetnost je normalno porazdeljena, s središčem v pričakovani globini. Verjetnost se izračuna tudi za barvno informacijo. Končna verjetnost je utežena vsota obeh verjetnosti.

Poleg ostalih značilk, bi lahko pri sledenju večih tarč upoštevali tudi njihovo medsebojno razdaljo. V [58] najdemo kombiniranje histogramov RGB, barvnih korelogramov in značilk LPB, ki so podlaga za izračun ujemanja detekcij.

Množenje

V primeru medsebojno neodvisnih značilk je možna njihova integracija z množenjem. Značilke modeliramo z verjetnostno funkcijo, ki modelira podobnost detekcij in obstoječih trajektorij. Verjetnostna funkcija je preprosto produkt verjetnosti posameznih značilk. Združevanje najdemo v [69], kjer so združene informacije o barvi in obliki, skupaj z množico drugih lokalnih značilk. Integracijo barvnih histogramov in detekcije ospredja najdemo v [56]. Produkt značilk oblike, teksture in globine pa sestavlja ogrodje v [18]. Pod to kategorijo pa lahko spada tudi scena, ki je razdeljena iz vidika večih kamer. Tu je vizualni model sestavljen iz barvnih značilk in ocene zasedenosti talne ravnine (angl. *ground occupancy estimation*).

Kaskadno združevanje

Gre za zaporedno obravnavo značilk z namenom, da se zmanjša preiskovalni prostor pri detekciji in kasnejšemu sledenju objektov. Primer najdemo v [17], kjer je zaporedno procesirana globina, zatem pa oblika in teksturne značilke. Z globinsko informacijo se generirajo hipoteze, zatem se aplicirajo ostale značilke. Vizualna informacija se obravnava tudi v [23], kjer se v prvi fazi detekcije povežejo v kratke sledi (glede na podobnost iz barvnega histograma, značilk HOG in deskriptorja optičnega toka). V drugi fazi pa se aplicira specifičen vizualni model (namenjen opisovanju pešcev), ki ga sestavlja več delov sledenega objekta.

4.2 Modeli gibanja

Modeli gibanja opisujejo dinamiko sledenih objektov. V primeru večih objektov je to še posebej pomembno, saj s predvidevanjem potencialnih pozicij objektov v prihodnosti lahko zmanjšamo prostor preiskovanja dejanskih pozicij. V splošnem se predvideva, da se

objekti gibljejo zvezno in brez sunkovitih sprememb smeri, čeprav so možni tudi posebni primeri izredno dinamičnega gibanja. Te primere je potrebno modelirati ločeno. Luo et al. modele gibanja razdeli v dve kategoriji: linearne in nelinearne modele [36].

4.2.1 Linearni modeli

Modelirani objekti se gibljejo s konstantno hitrostjo oz. konstantnim pospeškom. Model skoraj konstantne hitrosti (NCV) je Breitenstein et al. v [7] zapisal kot linearno kombinacijo:

$$(x, y)_t = (x, y)_{t-1} + (u, v)_{t-1} \cdot \Delta t + \epsilon_{x,y}, \quad (u, v)_t = (u, v)_{t-1} + \epsilon_{u,v}$$

Pri tem $(x, y)_t$ predstavlja napovedano pozicijo objekta v času t , $(u, v)_t$ pa njegovo novo hitrost v času t . Indeks $t - 1$ označuje vrednosti iz prejšnjega časovnega koraka, Δt pa velikost časovnega koraka. Vrednosti $\epsilon_{x,y}$ in $\epsilon_{u,v}$ sta naključni spremenljivki, porazdeljeni kot $\mathcal{N}(0, \sigma_{x,y}^2)$ in $\mathcal{N}(0, \sigma_{u,v}^2)$, ki modelirata negotovost stanja objekta. Varianca $\sigma_{x,y}^2$ sledi spremembam velikosti objekta, medtem ko je varianca $\sigma_{u,v}^2$ odvisna od uspešnosti sledenja (če je trajektorija blizu dejanski poti objekta, je varianca manjša, sicer se z napakami povečuje).

Avtorji Andriyenko in Schindler [3] ter Milan et al. [38] v svojih delih prav tako uporabljajo model konstantne hitrosti. Poleg tega uvedejo izraz za kaznovanje velikih razlik v hitrostih $C_{dyn} = \sum_{t=1}^{N-2} \sum_{i=1}^M \|v_i^t - v_i^{t+1}\|^2$, kjer je v_i^t hitrost tarče z indeksom i v času t . Izraz upošteva celotno zaporedje dolžine N sličic ter M tarč oz. trajektorij. V vsakem koraku se izračuna norma drugega reda nad soležnima vektorjema hitrosti. Z izračunanim faktorjem C_{dyn} se trajektorije ohranjajo zvezne oz. zglajene.

Xing et al. [65] v svojem modelu poleg trenutne hitrosti upošteva tudi nazaj usmerjeno hitrost v prejšnjem koraku. S tem lažje poveže lokalne sledi v trajektorije. Za primer vzemimo, da imamo dve sledi, T_i in T_j . Med njima nastane vrzel, ki sega od konca sledi T_i do začetka sledi T_j . Konec sledi T_i označimo z p_i , začetek sledi T_j pa z p_j . Gibanje naprej lahko modeliramo z normalno porazdelitvijo, ki je centrirana v $p_i + v_F$ z varianco Σ_i^F . Podobno se modelira tudi gibanje nazaj, ki ga opisuje normalna porazdelitev, centrirana v $p_j + v_B$ z varianco Σ_j^B . Končni model dobimo z množenjem porazdelitev gibanja naprej in gibanja nazaj:

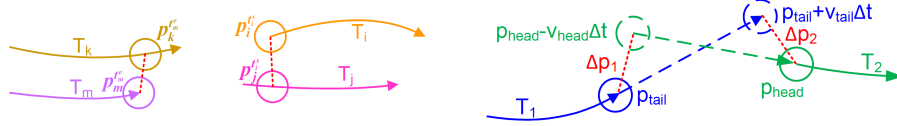
$$P_m(T_i, T_j) = \mathcal{N}(p_i + v_i^F \cdot \Delta t; p_j, \Sigma_j^B) \cdot \mathcal{N}(p_j + v_i^B \cdot \Delta t; p_i, \Sigma_i^F)$$

Yang in Nevatia [68] uvedeta drugačen pristop, kjer v enem vozlišču namesto ene de-

tekcije obravnavata lokalne sledi po parih (slika 4.1 levo). V ta namen uporabita pogojno naključna polja (CRF). Vsako vozlišče ima oznako (labelo), ki pove, ali sledi v paru lahko združimo v trajektorijo ali pa nakaže, da nista združljivi. Za labelo poskrbi enočlenski izraz v CRF modelu, ki pri izračunu upošteva vizualni model kot tudi informacijo o gibanju. Verjetnostna porazdelitev gibanja je določena kot odmik med ocenjenimi pozicijami linearnega modela in opazovanimi pozicijami iz detekcij. Ta model je predstavljen na sliki 4.1 desno, kjer imamo primer dveh lokalnih sledi T_1 in T_2 . Pri tem sta vpeljeni dve predpostavki. Prva je ta, da se T_1 izteče preden se pojavi T_2 . Druga predpostavka pa je ta, da med zaključkom T_1 in začetkom T_2 obstaja časovna vrzel z oznako Δt . Verjetnost povezovanja obeh sledi je odvisna od dveh pogojev. Prvi je odmik med ocenjeno in opazovano pozicijo zaključka T_1 , ki je definiran kot $\Delta p_1 = p_{head} - v_{head} \cdot \Delta t - p_{tail}$, drugi pa odmike med ocenjeno in opazovano pozicijo začetka T_2 , ki je definiran kot $\Delta p_2 = p_{tail} + v_{tail} \cdot \Delta t - p_{head}$. Verjetnostna porazdelitev gibanja v vrzeli je nato enaka

$$P_m(T_1, T_2) = \mathcal{N}(\Delta p_1; 0, \Sigma_p) \cdot \mathcal{N}(\Delta p_2; 0, \Sigma_p),$$

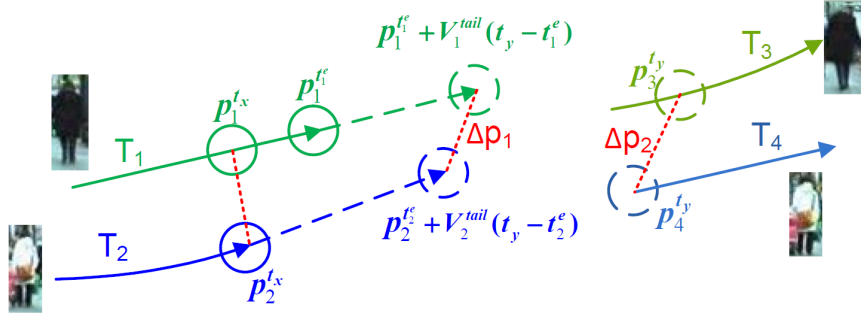
kjer je Σ_p kovariančna matrika negotovosti.



Slika 4.1 Primer parov bližnjih trajektorij (levo) in globalni model gibanja dveh trajektorij (desno) [68].

Opazimo lahko, da je ta model, ki je sicer precej popularen, zelo podoben predhodno opisanemu modelu [65]. Oba sicer obravnavata le en par sledi. Yang in Nevatia [68] sta zato dodala še model, ki analizira lokalne sledi zajete v dveh parih (slika 4.2), kar je nujno pri ločevanju večih objektov. V tem primeru za ilustracijo predpostavimo, da imamo štiri lokalne sledi, ki so paroma zapisane kot (T_1, T_3) in (T_2, T_4) . T_1 in T_2 predstavljata začetni sledi in sta si relativno blizu. Določi se čas $t_x = \min(t_1^e, t_2^e)$, ki označuje najbolj zgodnji čas, ob katerem sta še definirani končni točki T_1 in T_2 . Podobno se določi čas $t_y = \max(t_3^s, t_4^s)$, ki označuje čas, ko sta definirana začetka obeh sledi T_3 in T_4 . Za časovni znački t_x in t_y velja $t_x < t_y$. Zatam se oceni relativna razdalja med ocenama pozicij T_1 in T_2 v času t_y kot $\Delta p_1 = (p_1 + v_1^t \cdot (t_y - t_1^e)) - (p_2 + v_2^t \cdot (t_y - t_2^e))$, kjer sta v_1^t in v_2^t hitrosti objektov na zaključku trajektorij T_1 in T_2 . Dejansko relativno razdaljo lahko

preprosto določimo z izrazom $\Delta p_2 = p_3 - p_4$. Podobno kot v prejšnjem modelu, se tudi tukaj verjetnost povezovanja modelira z normalno porazdelitvijo $\mathcal{N}(\Delta p_1 - \Delta p_2; 0, \Sigma_p)$. Intuitivno si lahko model dveh parov razložimo na naslednji način: če pride do asociacije sledi T_1 s T_3 , potem obstaja velika verjetnost, da skupaj spadata tudi preostali dve sledi, v tem primeru T_2 in T_4 .



Slika 4.2 Modeliranje parov trajektorij iz [68].

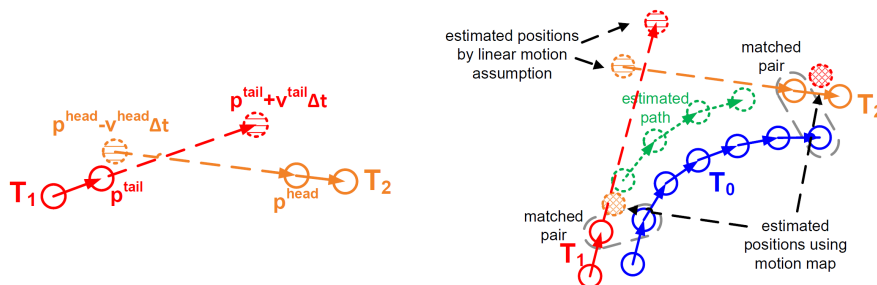
Kuo in Nevatia [31] poleg hitrosti upoštevata tudi pospešek. Pri tem definirata hitrost v časovnem koraku k kot $v_k = \frac{x_{k+1} - x_k}{t_{k+1} - t_k}$. Podobno definirata tudi pospešek koraka k kot $a_k = \frac{v_k - v_{k-1}}{0.5 \cdot (t_{k+1} - t_{k-1})}$. Verjetnost gibanja je zatem modelirana kot produkt normalno porazdeljenih spremenljivk, ki zajemajo pozicijo, hitrost in pospešek v vseh časovnih korakih k .

4.2.2 Nelinearni modeli

Linearni modeli so najpogostejše uporabljeni pri opisovanju gibanja. V določenih primerih kompleksnega gibanja so linearni opisi neustrezni, zato je potrebno vpeljati nelinearne modele, saj na ta način lahko natančneje določimo prave trajektorije gibanja.

Yang in Nevatia [70] nelinearni model uvedeta za rokovanje situacij, kjer prihaja do prostega oz. nepredvidljivega gibanja objektov. Kot vidimo na sliki 4.3, bi pri takšni postavitvi sledi T_1 in T_2 , linearni model iz [68] generiral nizko verjetnost povezanosti, čeprav T_1 in T_2 pripadata istemu objektu. Alternativen pristop je uvedba nelinearnega modela, ki je sestavljen iz množice vzorčnih sledi (T_0 na sliki 4.3 desno). V tem primeru sled T_0 razloži potek trajektorije, ki pripada vrzeli med T_1 in T_2 . Pravimo, da je T_0 podporna sled, saj obstajajo njeni elementi, ki se ujema tako z zaključkom T_1 kot tudi

z začetkom T_2 . Po postavitvi mostične trajektorije T_0 , se končna trajektorija sestavi podobno kot pri linearnih modelih.



Slika 4.3 Linearni model (levo), nelinearni model (desno) [70].

4.3 Modeli interakcije

Modeli interakcije zajemajo medsebojni vpliv sledenih objektov. Pri sledenju enega objekta so nepomembni, če pa obravnavamo več objektov so lahko zelo uporabno orodje. V zelo zgoščenem scenariju je namreč zelo verjetno, da bo posameznik podedoval lastnosti gibanja skupine, v kateri se nahaja. Konkretno, če opazujemo pot pešca lahko vidimo, da svojo hitrost in smer prilagaja tako, da se ne zaletava v mimoidoče. Podobne vzorce gibanja je moč prepoznati tudi v skupinah, saj se udeleženci med seboj usmerjajo in vodijo. Oba pojava sta dobro raziskana in modelirana kot socialni model [19] in model skupinskega gibanja [22].

4.3.1 Socialni modeli

V socialnih ali skupinskih modelih je predpostavljeno, da je posamezen objekt poleg drugih objektov odvisen tudi od vplivov okolja. V kolikor te vplive poznamo, si lahko z njimi pomagamo pri izboljšanju procesa sledenja teh objektov. Modeliranje je ločeno na dva vpliva: individualni vpliv (angl. *individual force*) in skupinski vpliv (angl. *group force*).

Individualni vpliv zajema ohranjanje destinacije (angl. *fidelity*) in ohranjanje smeri ter hkrati hitrosti gibanja (angl. *velocity constancy*). Skupinski vpliv po drugi strani zajema lastnosti celotne skupine. Lastnost privlačnosti (angl. *attraction*) pomeni, da člani skupine težijo k temu, da se med gibanjem nahajajo blizu drug drugega. Odbojnost

(angl. *repulsion*) zagotavlja, da kljub privlačnosti osebkki ohranjajo določeno minimalno medsebojno razdaljo. Koherentnost (angl. *conherence*) pa ohranja skupinsko gibanje na ta način, da se člani gibljejo s podobno smerjo in hitrostjo.

Pellegrini et al. [43] uvaja predpostavko, da vsak član skupine začne zelo zgodaj prilagajati svojo trajektorijo z namenom, da bi se tako izognil trkom z ostalimi udeleženci gibanja. Posameznik je modeliran s svojim stanjem (s_i) , ki ga sestavljata pozicija in hitrost v času t (p_i^t, v_i^t). Zatem se vzpostavi kriterij minimalne razdalje med vsemi člani skupine. Ta kriterij zagotavlja članom določen osebni prostor za manevriranje in izogibanje trkom ($D_{ij}(v_i)$). Soležno se definira pripadajoči energijski parameter (C_{ij}), ki je uporabljen v energijski enačbi interakcije $C_i^{inter}(v_i) = \sum_{i \neq j} w_{ij} C_{ij}(v_i)$ (skupaj z utežjo w_{ij}). V končni enačbi nastopata še parametra $C_i^{ssc}(v_i) = (u_i - \|v_i\|)^2$, ki kaznuje nenadno spremembo smeri in $C_i^{dfd}(v_i) = \frac{(\bar{p}_1 - p_1) \cdot v_1}{\|\bar{p}_1 - p_1\| \cdot \|v_1\|}$, ki kaznuje napako v doseganju destinacije (angl. *drift*). Končna oblika energijske funkcije je $C_i(v_i) = C_i^{inter}(v_i) + \lambda_1 C_i^{ssc}(v_i) + \lambda_2 C_i^{dfd}(v_i)$, kjer sta λ_1 in λ_2 parametra za uravnovežitev energijskih členov. Z optimizacijo energijske funkcije $C_i(v_i)$ se zmanjša kompleksnost preiskovalnega prostora za i -ti osebek, s tem pa je kasneje olajšan postopek iskanja njegove korespondence.

Yamaguchi et al. [66] prav tako predpostavlja, da se objekti obnašajo socialno. Destinacija vsakega objekta je določena ob upoštevanju t.i. osebnih, socialnih in okoljskih faktorjev, ki so formulirani kot členi v kriterijski funkciji. Člen dušenja (angl. *damping term*) kaznuje nenadne spremembe vektorja hitrosti, člen omejene hitrosti (angl. *speed term*) predstavlja ceno, ki je sorazmerna z odstopanjem od željene hitrosti, člen smeri (angl. *direction term*) kaznuje odstopanje od zastavljene destinacije, člen privlačnosti (angl. *attraction term*) predpostavlja, da tarče ostajajo skupaj, če se gibljejo v skupini, skupinski člen (angl. *group term*) kaznuje varianco hitrosti v skupini ter predhodno opisani člen interakcije, ki obravnava morebitne trke (iz [43]).

Qin in Shelton [48] sta prepričana, da upoštevanje socialnega obnašanja z uporabo gručenja izboljša identifikacijo in asociacijo objektov pri sledenju. Konkretno predpostavita, da vsi udeleženci na sceni spadajo v K skupin, kjer je K optimalno izbran parameter. Zahtevano je še, da je vsaka lokalna sled (ki jo naredi posameznik) konsistentna s povprečno trajektorijo skupine. Zaradi tega se pri asociaciji upošteva tudi razdalja med posamezno lokalno sledjo in omenjeno povprečno trajektorijo (v obliki dodatnega člena kriterijske funkcije).

Choi in Savarese [11] definirata dva socialna faktorja. Odbojni faktor skuša ločiti

tiste objekte, ki se potencialno nahajajo preblizu en drugega. Faktor skupinskega gibanja pa predpostavlja, da se relativna razdalja med dvema objektoma ohranja skozi čas (med zaporednimi sličicami). To hkrati pomeni, da tudi vektorja hitrosti obeh objektov ostajata podobna.

Scovanner in Tappen [52] za učenje dinamike pešcev v realnosti predstavita socialni model iz štirih členov. Člen C_{LM} omejuje gibanje tarče tako, da prepreči njeno preskakovanje v prostoru, kjer je prostor organiziran kot mreža polj. Člen C_{CV} ohranja konstantno hitrost tarč. Člen C_{Dest} vodi tarčo do njene destinacije. Člen C_{AV} pa za vsako tarčo upošteva ostale tarče, kar povzroči odbojno silo in tako prepreči morebitne trke med tarčami. Vsi členi uteženo nastopajo v energijski funkciji. Z optimizacijo funkcije se napove gibanje tarč, kar generira končne trajektorije.

Pellegrini et al. [42] hkrati rešuje problem asociacije in problem pripadnosti skupinam. Trajektorije se generirajo glede na gibanje in vizualno informacijo. Tarče v isti skupini pa morajo ohranjati medsebojno razdaljo in smer gibanja. Za reševanje in napoved se uporabi pogojno naključno polje (CRF) tretjega reda v kombinaciji z optimizacijo energijske funkcije.

4.3.2 Modeli gibanja množic

Vzorci gibanja iz področja simuliranja množic (pregled je na voljo s strani avtorja Zhan et al. [72]) lahko v sledenje vpeljemo v primeru, da sledimo individualnemu objektu znotraj večje množice. Tipično to velja za scenarije, ki so nasičeni z množicami objektov. Pri tem so objekti ponavadi dokaj majhni, zaradi česar pride do pomanjkanja oprijemljivih vizualnih značilk. V tem primeru je uporaba vzorcev gibanja v množici dober nadomestek vizualnim značilkam.

Ali in Shah [1] predpostavita, da je obnašanje posameznika odvisno od okoliških objektov in scenske postavitve (angl. *scene layout*). Za modeliranje tovrstnih vplivov so predlagata tri vrste talnih polj (angl. *floor field*). Statično polje (SFF) zajema strukturo scene in pri tem vključuje največkrat uporabljeno pot do scenskega izhoda. Dinamično polje (DFF) zajema gibanje množice okoli sledenega objekta. Mejno talno polje (BFF) pa upošteva vse fizične in virtualne ovire, ki se v sceni pojavijo.

Zhao et al. [74] sledi množici v strukturirani sceni z opazovanjem vzorcev gibanja. Te vzorce odkrijejo z metodo glasovanja N-dimenzionalnih tenzorjev (angl. *ND tensor voting*) avtorjev Mordohai in Medioni [41]. Sledi vzorcev gibanja so pridobljene s pomočjo

sledilnika KLT. Vzorci so predstavljeni kot 4-dimenzionalne točke, ki vsebujejo prostorski položaj in hitrost v dveh dimenzijah. S pomočjo vzorcev gibanja se nato oceni verjetnostna porazdelitev možnih detekcij, napove pa se tudi smer in hitrost objektov v posebnem položaju, kjer se sledenje izvaja.

Opazovanje skupine pešcev kaže na kolektivno strukturo gibanja v prostoru in času. Kratz in Nishino [25, 26] se gibanja naučita z množico skritih Markovih modelov (HMM). Celotna video sekvenca je predstavljena volumetrično. Lokalne regije so razdeljene na časovno-prostorske kocke. Vzorec gibanja v taki kocki je opisan s 3-dimenzionalno normalno porazdelitvijo, ki bazira na 3-dimenzionalnih gradientih vseh vokslov v tej kocki. Vzorec se spreminja v času glede na trenutno stanje, kar kaže na to, da ima Markovo lastnost. Gibanje v prihodnosti je zato lahko predvideno glede na prejšnja stanja (gibalne vzorce). Vzorec, ki je uporabljen za napovedovanje zato lahko generira omejitve, s katerimi se lahko izboljša proces sledenja.

Predhodno opisani modeli vzorcev predpostavljajo, da se skupine gibljejo koherentno in v skupni smeri. Predpostavka ne velja v primeru nestrukturiranega gibanja, kjer prihaja do različnih smeri gibanja. Za učenje različnih gibanj v sceni, Rodriguez et al. [50] predlaga koreliran tematski model (CTM). Objektom sledita dve uteženi komponenti. Prva je sledilnik, ki grobo oceni odmik vseh gibajočih se slikovnih elementov. Druga pa je sledenje vedenja na višjem nivoju. Poleg upoštevanja vizualne informacije se pridobijo tudi vzorci gibanja, katere se sistem predhodno nauči z nenadzorovanim učenjem iz velike baze posnetkov. Pri testnem videu se išče ujemanje prostorsko-časovnih zaplat (angl. *space-time patches*). Tako pridobljena informacija o gibanju kasneje nudi asistenco sledilniku, ki bazira na Kalmanovem filtru.

4.4 Modeli izključevanja

Model izključevanja je vključen v reševanje fizičnih trkov med objekti. Pri danih detekcijah in hipotezah obstoječih trajektorij se navadno upoštevata dve omejitvi [36]. Prva je izključevanje na nivoju detekcij, kar pomeni, da npr. dve različni detekciji v eni sličici ne moreta pripadati identični trajektoriji. Druga omejitev je izključevanje na nivoju trajektorij, kar pomeni, da si npr. dve trajektoriji ne moreta prilastiti identične detekcije. V nadaljevanju je predstavljeno modeliranje obeh omejitev.

4.4.1 Izključevanje na nivoju detekcije

Izključevanje na tem nivoju je modelirano s cenilnim izrazom, ki ga je predlagal Milan et al. [39]. Izraz kaznuje primere, ko sta dve hkratni detekciji (o_i^t in o_j^t v času t) pripisani k identični trajektoriji, čeprav sta med seboj dovolj oddaljeni, da te razdalje ni mogoče zanemariti.

Kumar K.C. in De Vleeschouwer [29] v sledenju v ta namen uporabita metodo propagiranja label (angl. *label propagation*). Izključevanje je modelirano s sestavljanjem grafa izključitev, ki je sestavljen iz detekcij. Omenjeni graf vsem detekcijam z enako časovno značko dodeli različne oznake oz. labele. Pri tem graf ni nujno popoln. Vsako vozlišče (ena detekcija) je povezano le s tistimi vozlišči, katerih detekcije so se zgodile v istem času. Uteži povezav so uniformne $w_{ij} = \frac{1}{n}$, kjer je n število detekcij, ki so se zgodile ob enakem času. Ko je graf zgrajen, se iz njega izračuna Laplaceova matrika L , napake oznak v grafu pa so predstavljene z optimizacijskim problemom, ki ga avtorja zapišeta kot $\arg \max_{Y \in S} \text{Tr}(YLY)$, kjer je Y vektor dodeljenih oznak za vsa povezana vozlišča, S je množica vrstično stohastičnih matrik, Tr pa je sled matrike (angl. *trace norm*) L .

Tematski model, ki bazira na Dirichletovem modelu mešanih procesov (DPMM) je predstavljen s strani avtorja Luo et al. [37]. Za modeliranje izključevanja uvede t.i. nemogoče povezave, ki posnemajo dejstvo, da dve prekrivajoči se lokalni sledi ne moreta pripadati identični gruči oz. trajektoriji.

4.4.2 Izključevanje na nivoju trajektorij

V poglavju linearnih modelov gibanja je bil že omenjen izraz za modeliranje dinamike v obliki člena energijske funkcije, ki ga uvedeta Andriyenko in Schindler [3]. Energijski člen iz tega prispevka, ki služi izključevanju pa je zapisan kot $C_{exc} \propto \sum_{t=1}^N \sum_{i=1}^N \frac{s_g}{\|p_i^t - p_j^t\|^2}$, kjer je s_g faktor skaliranja razdalje, vrednosti p_i^t in p_j^t pa sta talni koordinati objektov i in j . V primeru, da sta si objekta zelo blizu, to pomeni, da bo vrednost člena C_{exc} zelo visoka, kar vpliva na večjo ceno energijske funkcije. V svojem drugem delu Andriyenko et al. [2] izključevanje trajektorij prav tako rešuje z uporabo cenilnega izraza. Tukaj je cena kazni inverzno proporcionalna minimalni razdalji trajektorij, ki se časovno prekrivata. V tem primeru se ena od trajektorij opusti oz. zavrže, kar prepreči kolizijo.

Milan et al. [39] sankcionira dve trajektoriji, ki sta si blizu, vendar imata različni labeli. Kazen je (podobno kot prej) tudi tukaj proporcionalna časovno-prostorskemu prekrivanju med obema trajektorijama (bližje sta si, večja je kazen).

Butt in Colins [10] izključevanje zajameta kot dodatno omejitev v objektivni funkciji omrežnega pretoka. Iz množice detekcij v dveh zaporednih sličicah se zgradi graf, kjer je vsako vozlišče par ujemajočih se detekcij, vsaka povezava pa predstavlja del omrežnega pretoka. Povezava se lahko nahaja v dveh stanjih. Če ima njen pretok vrednost 1, to pomeni, da sta njeni vozlišči povezljivi. V nasprotnem primeru ima pretok vrednost 0, kar označuje, da vozlišči nista povezljivi. Če upoštevamo predpostavko, da ena detekcija lahko pripada le eni trajektoriji, potem to pomeni, da je omejitev pretoka (in s tem omejitev izključevanja) vseh povezav omejena na 1.

4.5 Reševanje prekrivanja

Prekrivanje objektov je kritičen problem v sledenju. Povzroči lahko zamenjavo identitet ali fragmentacijo trajektorij. Poznamo več tipov prekrivanja. V primeru, da sledimo gibalno dinamičen objekt (npr. osebo), lahko ta zaradi svojega gibanja prekrije del sebe. Pri večih objektih lahko prihaja do medsebojnega prekrivanja, še posebej kadar pride do njihove interakcije. Prekrivanje pa lahko povzroči tudi scena oz. ozadje, ki lahko delno ali popolno prekrije objekt, ki mu sledimo. Obstaja več strategij za odpravljanje teh težav (kategorizacija iz [36]).

4.5.1 Del proti celoti

Tip prekrivanja “del proti celoti” izkorišča predpostavko, da ob delnem prekrivanju del objekta ostaja viden. Glede na vidni del objekta zatem sklepamo, kje se objekt točno nahaja. V primeru popolnega prekrivanja še vedno velja, da del objekta vidimo, preden se popolnoma prekrije – tako lahko vsaj ocenimo smer premikanja in sledilnik usmerimo proti oceni. Pri tej strategiji se zato uporabljajo sestavljeni vizualni modeli.

Hu et al. [21] predlaga bločno razdeljeni model (angl. *block-division model*), ki poleg prekrivanja skuša reševati tudi obnovitev stanja objekta po prekrivanju. Objekt v tem modelu je razdeljen na več blokov, ki se med seboj ne prekrivajo. Za vsak blok se na podlagi učenja podprostorov (angl. *subspace learning*) konstruira vizualni model. Funkcija podobnosti (angl. *likelihood*) se določi glede na velikost napake pri rekonstrukciji objekta med prekrivanjem v vsakem bloku posebej. Prednost razdeljenega modela je ta, da je prostorska informacija zajeta v produktu funkcij podobnosti vseh blokov. Poleg tega se lahko iz napak rekonstrukcije sestavi zemljevid prekrivanja (angl. *occlusion map*), ki je lahko v nadaljevanju uporabljen za obrazložitev relacij med vpletenimi objekti. Konkre-

tno to pomeni, da se lahko blokovne vizualne modele selektivno osvežuje, kar izboljša robustnost postopka.

Yang in Nevatia [67] uporabita večdelni model za razločevanje objektov v ospredju in v ozadju. Natančneje objekt razdelita na 15 delov. Vizualni model se navezuje na lokalno sled in je sestavljen iz množice značilk in pripadajočih uteži. Verjetnost povezljivosti dveh lokalnih sledi se določi kot podobnost med začetkom ene in koncem druge sledi. Pomen razdeljenega modela pri tem je, da v primeru prekrivanja ne upošteva značilk pridobljenih iz prekritih delov objekta. Učenje vizualnega modela poteka preko ojačitvenega učenja (angl. *boosting*).

Izadinia et al. [23] razdeljeni model uvede pri sledenju večih oseb. Za detektor uporabi večdelni detektor oseb, ki je delo avtorja Felzenszwalb et al. [15]. Detekcija in sledenje se izvajata na dveh nivojih – na nivoju celotnega telesa in obenem na nivoju posameznih okončin. Ocena končne trajektorije je tako pridobljena s skupno asociacijo celega telesa s posameznimi deli. V primeru prekrivanja zagotovo prihaja do težav z detekcijo na nivoju celotnega telesa. V tem primeru težave rešujejo detekcije vidnih okončin, preko katerih se najprej ocenijo trajektorije na njihovem nivoju, zatem pa se rekonstruira še končna trajektorija, ki je sestavljena iz obeh nivojev. Podoben model je predlagan tudi s strani avtorja Shu et al. [54].

V splošnem velja, da sledenje na podlagi vizualnih modelov med prekrivanjem lahko odpove. Podobno kot pri razdeljenih modelih, kjer lahko še vedno opazujemo določene vidne dele med prekrivanjem, to velja tudi za metode sledenja z gručenjem značilnih točk (angl. *feature point clustering based tracking*). Tukaj velja predpostavka, da točke s skupnimi lastnostmi gibanja pripadajo enakemu objektu. Na ta način – podobno kot pri vizualnih razdeljenih modelih – lahko rešujemo pojav prekrivanja. Določanje trajektorij iz točk deluje, dokler obstajajo vidni deli objekta.

4.5.2 Testiranje hipotez

Testiranje hipotez je strategija, ki pri reševanju prekrivanja uvaja hipoteze, ki jih testira na trenutno dostopnih opazovanjih.

Zhang et al. [73] predlaga eksplicitni model prekrivanja (EOM). Model je integriran v ogrodje za reševanje dolgoročnih prekrivanj in določanja korespondence oz. asociacije podatkov. V tipičnem primeru korespondenca podatkov pravi, da sta dve sledi povezljivi le, če je vrzel med njima dovolj majhna. Ta kriterij sicer lahko pripelje do fragmentacije v

končni rešitvi. Če se po drugi strani velikostni prag vrzeli zviša, se s tem nekoliko olajša povezljivost sledi, vendar se s tem poveča verjetnost napačnih asociacij. Da bi rešili ta problem, avtorji generirajo hipoteze možnih prekrivanj, ki so osnovane na določenih omejitvah. Če je npr. razlika v velikosti ali razdalji dveh opazovanj primerno majhna, potem velja, da sta vpletena objekta možna kandidata za prekrivanje. Hipoteze skupaj z obstoječimi sledmi predstavljajo vhod v ogrodje, kjer se v nadaljevanju s pomočjo metod MAP določi optimalna rešitev.

Podoben model, ki prav tako bazira na hipotezah in njihovem testiranju, v svojem delu vpelje Tang et al. [59]. V nasprotju z ostalimi pristopi, kjer je prekrivanje predstavljeno kot moteč dejavnik, se tukaj obravnava kot podpora detektorju, saj opazovanje prekrivanja odkriva določene vizualne vzorce. V ta namen se zgradi dvojni detektor oseb, s katerim se zajamejo različni nivoji prekrivanja dveh oseb. Detektor se uči iz sintetično združenih parov objektov z različnimi nivoji prekrivanja, kar pomeni, da vnaprej pozna faze prekrivanja. Slika 4.4 prikazuje primere sintetiziranih učnih vzorcev, iz katerih se dvojni detektor uči. V kombinaciji z detektorjem klasičnega tipa lahko predstavlja temelj sledenja večjega števila prekrivajočih se tarč.



Slika 4.4 Primeri sintetiziranih prekrivanj (delež prekrivanja narašča od leve proti desni) [59].

4.5.3 Predpomnjenje in obnovitev

Ta strategija zajema opazovanja, ko pride do prekrivanja in si zapomni stanja objektov preden se prekrijejo. Ko se prekrivanje konča se stanje objektov obnovi glede na predhodno shranjena stanja pred prekrivanjem.

Mitzel et al. [40] združi dva sledilnika, od katerih eden bazira na segmentaciji slike, drugi pa na visoko nivojski detekciji objektov. Pri tem visoko nivojska detekcija skrbi za odkrivanje novih trajektorij, sledilnik s segmentacijo pa rešuje lokalno asociacijo v zaporednih sličinah. Ko pride do prekrivanja, lokalno sledenje s segmentacijo navadno odpove. Zaradi tega visoko nivojski sledilnik ohranja trajektorijo aktivno še do 15 sličic po dogodku in hkrati razširja (ekstrapolira) pozicijo mirujoče trajektorije v tem časovnem

okviru. V primeru, da se objekt ponovno pojavi, se trajektorija aktivira in identiteta objekta se ohrani.

Ryoo in Aggarwal [51] predlagata t.i. strategijo “opazuj in razloži” (angl. “*observe-and-explain*” strategy) za reševanje prekrivanja med objekti ter tudi prekrivanja objektov s samo sceno. Strategija lahko zmanjša računsko kompleksnost sledenja, ker se način opazovanja aktivira le v primeru, ko pride do negotovosti med dogodkom prekrivanja. Ko se nabere dovolj evidence iz opazovanj, se generirajo rešitve v obliki t.i. korespondenčnih razlag opazovanj.

4.5.4 Ostalo

Predhodno opisane strategije ne pokrivajo celotnega področja, ki se ukvarja z reševanjem prekrivanja med sledenjem. V praksi obstaja preprostejša metoda, ki prekrivanje obravnava kot presek očrtanih pravokotnikov, vendar ne deluje v vseh primerih [36]. Velikokrat pa najdemo tudi različne kombinacije predhodno opisanih strategij.

4.6 Primeri verjetnostnih metod

V nadaljevanju na kratko opišemo znane verjetnostne metode (rekurzivne Bayesove filtre, filtre z delci, filtre JPDA in metodo MHT). Metode na različne načine rešujejo opisano problematiko, so različno kompleksne in imajo svoje prednosti in slabosti.

4.6.1 Rekurzivni Bayesovi filtri

Najbolj prepoznavna metoda iz te kategorije je uveljavljeni Kalmanov filter (angl. *Kalman filter*) [45]. Metoda je zasnovana kot sistem enačb, s katerim lahko predvidimo potek opazovanega sistema v katerem nastopa določena negotovost. Beseda filter v imenu metode se nanaša na dejstvo, da se z ocenjevanjem vrednosti opazovanega parametra iz meritev odstrani tudi prisotna negotovost oz. šum.

Osnovna različica Kalmanovega filtra predpostavlja linearno dinamiko in normalno porazdeljeno negotovost sistema. V primeru, da ta dva pogoja držita, Kalmanov filter predstavlja optimalen način ocenjevanja vrednosti parametrov, ker v osnovi minimizira srednjo kvadratno napako teh parametrov. V nasprotnem primeru osnovna različica filtra ne poda optimalne rešitve.

Da lahko Kalmanov filter uporabimo tudi v nelinearnih sistemih so potrebne prilagoditve. Obstaja veliko razširitev in dopolnitev osnovne različice. Ena izmed izboljšav

je razširjeni Kalmanov filter (EKF) [45], ki deluje povsem enako kot osnovna izvedba filtra, s to razliko, da se izvede predhodna linearizacija uporabljenega dinamičnega modela. Slabost razširitve EKF je v tem, da se v določenih primerih kovarianca negotovosti napačno propagira v rekurzivni shemi filtra. Obstaja tudi različica UKF [45], pri kateri se dinamični model aproksimira s kontrolnimi točkami in numerično integracijo.

Vse tukaj našteje različice ne rešujejo problema večih hipotez, zato so bile predlagane tudi druge različice združenih Kalmanovih filtrov (angl. *multiple Kalman filters*) oz. metode iz področja mešanih Gaussovih porazdelitev (angl. *mixture of Gaussians*) [45].

4.6.2 Filtri z delci

Filtri z delci (angl. *particle filters*) so posplošitev Kalmanovega filtra [45]. Osnovo predstavlja kondenzacijski algoritem (angl. *condensation algorithm*), ki je združen s principom Monte Carlo integracije (angl. *Monte Carlo integration*). Ta princip pravi, da je poljubno verjetnostno porazdelitev mogoče aproksimirati z dovolj velikim številom vzorcev, ki jih izvlečemo iz te distribucije. Uporaba teh filtrov je pogosta v primerih, ko obnašanja oz. negotovosti objektov ne moremo opisati z linearnimi modeli in normalno porazdelitvijo. Namesto analitičnih modelov se zato z uporabo kondenzacijskega algoritma uvede faza vzorčenja (angl. *sampling*), s katero se aproksimira problematična multimodalna porazdeljenost stanja objektov. Primeri uporabe v [7, 21, 27, 65].

4.6.3 Filtri JPDA

Obstaja več različic filtrov JPDA. Primer najdemo v [60], kjer je predstavljen t.i. filter Monte Carlo JPDA, ki je primeren za nelinearne probleme s kompleksnimi verjetnostnimi modeli. Glavna ideja metode je rekurzivno prilagajanje marginalne distribucije za vsak sleden objekt. Izračun teh distribucij se izvede z uporabo kondenzacijskega algoritma. Zaradi negotovega izvora tarč, izračuna teh distribucij ni mogoče opraviti neodvisno, kar pomeni, da je potreben dodaten postopek za reševanje podatkovne asociacije, kar pomeni prireditve novih položajev obstoječim tarčam.

4.6.4 Metoda MHT

Metoda sledenja večim hipotezam (MHT) je izjemno zmogljiva, vendar hkrati kompleksna metoda za sledenje večih objektov z generiranjem in ohranjanjem majhnega seznama potencialno ustreznih hipotez. Ključna strategija je v tem, da se asociacijske odločitve

zakasnijo do te mere, dokler ni možna razrešitev vseh nejasnosti, ki so s tem povezane. Metoda sestoji iz konstruiranja in osveževanja dreves s hipotezami, postopka aktivacije (angl. *gating*), ocenjevanja trajektorij, formiranja globalnih hipotez in rezanja odvečnih hipotez iz predhodno omenjenih dreves (angl. *prunning*). Posamezno drevo je na začetku zgrajeno iz vseh možnih hipotez, ki izvirajo iz enega opazovanja. Predhodno obstoječa drevesa so prav tako posodobljena s hipotezami iz trenutnih opazovanj. Vsaka hipoteza v drevesu je razširjena z opazovanji, ki se nahajajo v aktivacijskem območju te hipoteze. Aktivacijsko območje je določeno glede na oceno gibanja in označuje pričakovano območje opazovanj v prihodnosti. Kandidati trajektorij zatem dobijo svojo oceno (angl. *score*), ki je sestavljena kot utežna vsota ocene gibanja in vizualne ocene. Globalne trajektorije se zatem določijo z uporabo metode MWIS (ki je opisana v nadaljevanju). Zaradi eksponentnega naraščanja števila novih hipotez pa se opravi še rezanje dreves, ki odstrani neveljavne hipoteze in prazna drevesa. Podrobnejši opis metode je na voljo v [24].

4.7 Primeri determinističnih optimizacij

Tukaj navajamo še nekaj determinističnih metod, ki se v literaturi pogosto pojavljajo (ujemanje dvodelnih grafov, dinamično programiranje, maksimalno pretočno omrežje z minimalno ceno, metoda CRF ter metoda MWIS).

4.7.1 Ujemanje dvodelnih grafov

Dve disjunktni množici vozlišč v grafu lahko predstavljata obstoječe trajektorije in nove detekcije (v zaporednem oz. koračnem sledenju) ali pa dve množici lokalnih sledi (v primeru bločnega sledenja). Uteži med vozlišči so modelirane kot podobnosti (afinitete) med trajektorijami in detekcijami. Zatem se za določitev ujemanja vozlišč iz obeh množic uporabi požrešna metoda dodelitve (angl. *greedy bipartite assignment algorithm*) ali pa metoda iskanja globalnih sosedov. Pristop najdemo v [7, 48, 54, 65].

4.7.2 Dinamično programiranje

Dinamično programiranje je sistematičen pristop, ki preveri vse možne poti pri reševanju določenega problema. Ideja reševanja temelji na optimalni podstrukturi problemov, kar pomeni, da je vsak del optimalne rešitve prav tako optimalen. Na področju sledenja se uporablja predvsem za reševanje problema asociativnosti. V literaturi zasledimo več

različic realizacije, npr. linearno programiranje (angl. *linear programming*), logično kvadratično programiranje (angl. *boolean quadratic programming*), iskanje K-najkrajših poti (angl. *K-shortest path routing*). Primer pristopa v obliki iskanja najkrajših poti najdemo v [4].

4.7.3 Maksimalno pretočno omrežje z minimalno ceno

Metoda se imenuje tudi transportno omrežje (angl. *Min-cost max-flow network flow*). Osnovna ideja je usmerjen graf, kjer ima vsaka povezava določeno kapaciteto. Na področju sledenja objektov, so običajno meritve (osnovne detekcije ali delne sledi) predstavljene z vozlišči, pretok (povezave med vozlišči) pa je modeliran kot indikator, ki pove, ali dve vozlišči spadata skupaj (vrednost indikatorja postane 1) ali ne (vrednost indikatorja postane 0). Da se zadosti zahtevi po uravnoteženosti pretoka se v graf doda tudi izvirno (začetek trajektorije) in zaključno vozlišče (konec trajektorije). Ena pretočna pot v grafu tako predstavlja eno trajektorijo. Število pretočnih poti iz izvirnega vozlišča v zaključno vozlišče je enako številu objektov, ki jim sledimo (število trajektorij). Cena celotnih pretočnih tranzicij pa je predstavljena kot verjetnostna porazdelitev vseh asociacijskih hipotez. Različice pristopa med drugim najdemo v [10, 63, 73].

4.7.4 Pogojno naključna polja

Tudi metoda pogojno naključnih polj (CRF) definira graf, kjer pa meritve predstavljajo vhodne podatke. Vsako vozlišče v grafu predstavlja par delnih sledi. Vozlišču pripada oznaka (labela), ki pove, ali je par v vozlišču možno združiti (vrednost labele postane 1) ali ne (vrednost labele postane 0). Z množico label se zatem sestavi zemljevid korespondenčnih sledi, iz katerega se določi optimalna rešitev. Opazimo lahko podobnost ideje v primerjavi s predhodno omenjenim pristopom pretočnega omrežja. Uporabo najdemo v [39, 68].

4.7.5 Neodvisna množica z maksimalno težo

Metoda MWIS je še ena metoda, ki v osnovi uporablja ogrodje v obliki grafa. Vozlišče tudi tukaj predstavlja par sledi v zaporednih sličicah, poleg tega pa pripadajoče uteži vozlišč definirajo podobnost omenjenih sledi. Povezava dveh vozlišč je omogočena v primeru, da imajo vpletene sledi skupne detekcije. Z uporabo tovrstnega sestavljenega grafa se ponavadi rešuje problem podatkovne asociacije (primer najdemo v [8, 24]).

5 Praktični del

Praktični del začnemo s študijo primerov uporabe, ki zajema uporabo svetlobnih oznak in kamer ter stanje okolja, v katerem naj bi sistem deloval. Zatem sledi opis uporabljene strojne opreme, nazadnje pa so v programskem delu opisane tudi podrobnosti o implementaciji sistema.

5.1 Študija primerov uporabe

V tem delu naloge predstavimo zahteve in razmere okolja, v katerem naj bi sistem deloval. Opišemo vlogo svetlobnih oznak, naštejemo možne osvetlitve prizorišča ter omenimo možne postavitve kamer.

5.1.1 Primeri uporabe svetlobnih oznak

Predvideli smo dve možni uporabi svetlobnih oznak. Prva je uporaba ene same oznake, ki jo oseba pritrdi na sredino prsnega koša (slika 5.1). Druga pa je uporaba večjega števila oznak, ki so razporejene po kompletu naglavnega mikrofona (slika 5.2). Kot bomo videli v nadaljevanju je uporaba prve možnosti bolj smiselna, saj tako lažje modeliramo objekte

v sledilniku spektra IR.



Slika 5.1 Prva improvizirana različica oznak IR, ki je pritrjena na prsnem košu (levo pogled v vidnem spektru, desno pogled v spektru IR).



Slika 5.2 Druga različica združenih oznak IR (levo pogled v vidnem spektru in desno pogled v spektru IR).

5.1.2 Primeri možne osvetlitve prizorišča

Sledilni sistem naj bi bil uporabljen v dvoranskih prostorih, kjer bi deloval kot podpora za lokalizacijo oseb. Pri tem smo predpostavili štiri kategorije osvetljenosti prizorišča:

- osvetljen oder, zatemnjena sedišča (predstava, prireditev),
- zatemnjen oder, osvetljena sedišča (predavanje, projekcija),
- zatemnjena dvorana (samo projekcija) in
- osvetljena dvorana (drugi dogodki, brez uporabe osvetlitve oz. zatemnjenosti).

5.1.3 Primeri možne postavitve kamer

Predvidoma sta možni dve postavitvi kamer. Postavitev nad prizoriščem je po našem mnenju slabša izbira, saj je postavitev zahtevnejša, možnosti naknadnih popravkov kalibracije kamer so omejene, poleg tega pa bi za to postavitev potrebovali nadgradnjo detektorja oseb v vidnem spektru. Zaradi naštetih težav smo se odločili za lažjo možnost – to je postavitev kamer pred prizoriščem oz. med sedišči. V tem primeru omenjeni detektor deluje brez težav, hkrati pa so kamere na dosegu roke, zaradi česar je olajšano kalibriranje. Možnost postavitve kamer nad prizoriščem vseeno puščamo odprto, saj bi na ta način lahko bolje reševali interakcijo in trke med objekti.

5.2 Strojna oprema

V nadaljevanju opišemo katera dodatna strojna oprema je bila uporabljena pri realizaciji sledilnega sistema. Predstavljeni so modeli kamer, ki so bili uporabljeni tekom razvoja in testiranja. Opišemo tudi svetlobne oznake, ki so uporabljene v navezi s sledilnikom spektra IR.

5.2.1 Kamere

Podatke smo zaradi implementacijskih razlogov (hitrosti) ter zmogljivosti kamer zajemali pri manjših ločljivostih (320×240). Ker sledilnika delujeta ločeno, vsak zajema video tok iz svoje kamere. To pomeni, da imamo 2 kamere, ki sta trenutno ročno kalibrirani. Pri tem ima ena kamera namesto filtra, ki blokira infrardečo svetlobo, namensko plast za blokiranje vidne svetlobe. V nadaljevanju opisujemo dva modela kamer (slika 5.3), ki smo jih preizkusili v ta namen.

Logitech QuickCam 4000 spletna kamera je starejša spletna kamera z največjo možno ločljivostjo 320×240 ter vmesnikom USB. Osveževanje znaša od 5 do 25 sličic/s, kar predstavlja težave, zlasti pri dinamičnih svetlobnih pogojih, saj se senзитivnost senzorja prilagaja trenutni osvetlitvi. Posledica tega je variabilno število sličic na sekundo, česar pa si pri sledenju tarč ne želimo. Kljub slabostim je prednost kamere izredno dobra senзитivnost, saj smo svetlobne oznake lahko sledili tudi na večjih razdaljah (+25m). Zaradi te prednosti smo sledenje v veliki večini primerov izvajali s kamerama tega tipa.

Sony PS3 Eye kamera je nizkocenovna USB kamera, primarno namenjena za uporabo skupaj z igralno konzolo Sony PS3. Kamera je sicer namensko izdelana za delovanje v pogojih z nizko ambientalno svetlobo, kar je eden izmed glavnih razlogov za uporabo te kamere v našem primeru. Hkrati ponuja zajemanje posnetkov s hitrostjo do 120 sličic/s pri ločljivosti 320×240 (oz. 60 sličic/s pri ločljivosti 640×480). Glede na to, da je v našem primeru odzivnost osveževanja pomembnejša od višje ločljivosti, ter da kamera izredno dobro deluje tudi v dinamičnih svetlobnih pogojih, se je izkazala za boljšo alternativo predhodno opisani Logitechovi spletni kameri. Edini problem (ki se je pokazal šele med testiranjem) je ta, da kamera pri zajemanju posnetkov na večjih razdaljah (+10 m) izgublja senzitivnost, kar posledično pomeni, da se zaradi tega začnejo izgubljati tudi svetlobne oznake, ki jih uporabljamo med sledenjem. Zaradi te glavne pomanjkljivosti kameri tega tipa nista bili velikokrat uporabljeni. V morebitni produkcijski fazi bo torej potrebno poiskati kamere z nekoliko primernejšimi specifikacijami.



Slika 5.3 Preizkusna modela kamer v fazi testiranja: Sony PS3 Eye (levo) in QuickCam 4000 (desno).

5.2.2 Oznake IR

Pri oznakah je bil pomemben predvsem vidik čim manj invazivnega pritrjevanja in nošenja ter omejitev velikosti. Za infrardeče svetlobne oznake smo v našem primeru uporabili generične infrardeče led diode z valovno dolžino med 940 nm in 950 nm. Ker so diode relativno majhen porabnik, so bili za napajanje uporabljeni baterijski vložki z oznako CR2032.

Uporabljene generične diode ustvarjajo usmerjen (reflektorski) žarek svetlobe, kar pomeni, da se snop svetlobe fokusira v območju, ki zajema od 0° do pribl. 15° kota, ki je merjen frontalno od vrha diode. To v praksi pomeni, da diode niso enako vidne v različnih orientacijah oz. položajih. Med sledenjem lahko usmerjenost žarka oz. njegova odsotnost predstavlja težavo, saj ga kamera lahko ne vidi ali pa premočan žarek kamero zaslepi. Da bi izboljšali difuzne karakteristike diod, smo diode predhodno obdelali z brusnim papirjem tipa P60. Difuznost smo še dodatno izboljšali z dodajanjem ovoja iz vročega lepila neposredno na svetilno površino vsake diode.

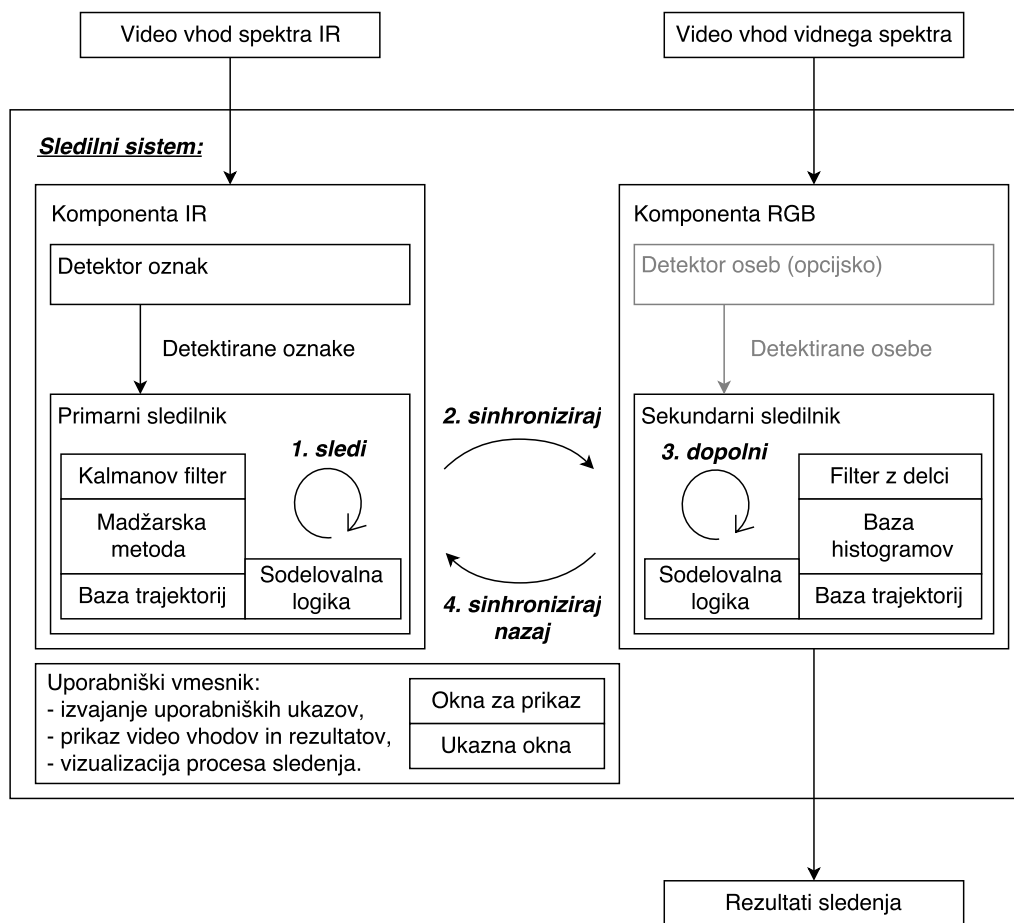
5.3 Programski del

Sistem je realiziran v ogrodju `OpenFrameworks`, ki združuje več pogosto uporabljenih knjižnic, zaradi česar se razvijalcu poenostavi proces integracije različnih programskih tehnologij. Poleg tega ogrodja je bila uporabljena tudi knjižnica `dlib`, ki nudi podporo matričnim operacijam ter paralelna implementacija gradientnega detektorja oseb `fastHOG`, ki je bila uporabljena v sledilniku vidnega spektra.

Diagram celotnega sistema najdemo na sliki 5.4, kjer vidimo, da sistem sestavljajo tri glavne komponente: primarni sledilnik (v spektru IR), sekundarni sledilnik (v vidnem spektru) in uporabniški vmesnik. Shema delovanja obeh sledilnikov je v osnovi podobna. Na nivoju sledenja oba obravnavata objekte individualno (vsakega posebej), kar vključuje model opazovanja in dinamični model. Na nivoju ohranjanja identitet, interakcije in asociacije pa primarni sledilnik uporablja Madžarsko metodo [28] (implementacija metode je iz knjižnice `dlib`). Sekundarni sledilnik pa za reševanje problema identifikacije uporabi barvno informacijo v obliki histogramov. Oba sledilnika sicer vsebujeta še nekatere specifične dodatne predpostavke, ki so bile zastavljene med načrtovanjem sistema. V nadaljevanju opisujemo podrobnosti obeh sledilnikov, njuno delovanje in uporabljene predpostavke ter tudi morebitne razlike. Opišemo pa tudi večnitnost sistema, medsebojno sodelovanje obeh sledilnikov in tretjo komponento sistema – uporabniški vmesnik.

5.3.1 Sledilnik v spektru IR

Detekcija objektov v primarnem sledilniku je zaradi uvedbe svetlobnih oznak relativno preprosta in se izvaja zelo hitro. Sledenje je realizirano s Kalmanovim filtrom, za ohranjanje identifikacije pa (podobno kot v [49]) skrbi Madžarska metoda. Poleg tega sledilnik



Slika 5.4 Shematski diagram komponent, ki so vključene v sledilni sistem z osnovno zanko procesiranja.

vsebuje sinhronizacijsko logiko za sodelovanje s sekundarnim sledilnikom in začasni pomnilnik, v katerem se hrani nedavna zgodovina trajektorij.

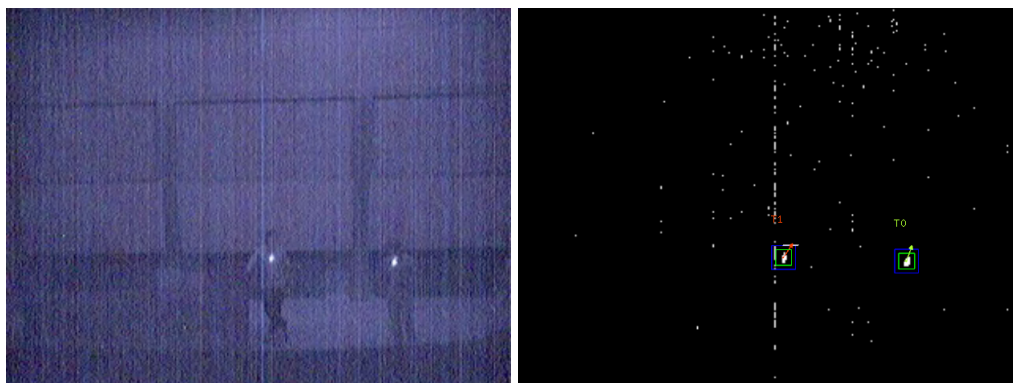
Detektor

Detektor primarnega sledilnika je zaradi strojne modifikacije kamere zelo enostaven. Rezultat opravljene modifikacije je ta, da senzor v kameri v idealnih razmerah zajame le infrardečo svetlobo, ki prihaja iz oznak IR. Idealne razmere v našem scenariju so popolna odsotnost drugih infrardečih svetlobnih virov ter primerna oddaljenost objektov od kamere. Prvi pogoj olajša programski del implementacije, saj se odstrani zahtevo po filtriranju detekcij, ki ne predstavljajo oznak (angl. *false detections*). Drugi pogoj obstaja le zaradi strojnih omejitev kamer, ki so bile uporabljene v razvoju. Pri večini

nizkocenovnih kamer (predvsem pri senzorjih tipa CMOS) se namreč izkaže, da je senzor pri nizki globalni osvetljenosti premalo občutljiv na zelo oddaljene infrardeče svetlobne vire, kar pomeni, da jih od določene razdalje naprej ne zazna več.

V programskem delu detektorja se nato opravi le še navadno (enakomerno) upravljanje in iskanje regij. Obe opravili se opravita s standardnimi metodami iz knjižnice `OpenCV`. Najdene regije (angl. *blobs*) zatem omejimo glede na velikost in podolgovatost. Centralni momenti regij, ki ustrezajo vsem kriterijem zatem predstavljajo detekcije, ki jih na vhod sprejme sledilnik.

Trenutno ena detekcija ustreza enemu objektu (primer na sliki 5.5). V kolikor bi želeli uporabiti drugi tip oznak IR - torej na enem objektu uporabiti več tovrstnih oznak, bi potrebovali dodatno logiko, ki bi bila sposobna rekonstruirati prostorski položaj trenutno vidnih oznak in jih hkrati dodeliti pravemu objektu, še posebej v primeru prekrivanja ali neposredne bližine dveh objektov. Reševanje tega problema je zahtevno in obsežno, zato ga bomo v naši nalogi izpustili – lahko pa predstavlja dobro osnovo za izboljšave in nadaljnje delo.



Slika 5.5 Primeri vhodne slike v detektor (levo) in rezultat detektorja po upravljanju – s pripadajočo aktivacijo trajektorij T_0 in T_1 (desno).

Sledilnik

Sledilnik rešuje nekoliko več problemov kot detektor. Poleg določanja korespondence med novimi detekcijami in obstoječimi trajektorijami, se sooča tudi s procesom pojavljanja detekcij, ki tvorijo nove trajektorije ter s problemom odstranjevanja zastarelih trajektorij. Pri vsem tem je najbolj pomembno ohranjanje konsistentnosti identitet, ki so dodeljene

trajektorijam.

Nove lokacije aktivnih trajektorij ocenimo z uporabo Kalmanovega filtra [45]. Za dinamični model uporabimo model skoraj konstantne hitrosti (NCV) [7]. Filter v vsakem koraku sledenja napove nove pozicije oznak, posodobi negotovosti svojih napovedi in v bazo predikcij vključi pozicije novih detekcij. Detekcije zatem priredimo aktivnim trajektorijam z Madžarsko metodo [28]. Detekcije z manjkajočo prireditvijo aktivirajo nove trajektorije.

Določanje korespondence trajektorij in detekcij: Ohranjanje identitet se izvede z Madžarsko metodo. Metoda je namenjena reševanju linearnega problema dodeljevanja (angl. *linear assignment problem*). Vhod v metodo je cenilna matrika (angl. *cost matrix*) C , ki jo sestavimo z upoštevanjem položaja novih detekcij in obstoječih (aktivnih) trajektorij v vsakem koraku sledenja. Število vrstic matrike C je enako številu aktivnih trajektorij (m), število stolpcev matrike C pa je enako številu novih detekcij (n). Vsaka celica matrike ($c_{i,j}$) predstavlja ceno med i -to trajektorijo in j -to detekcijo in je definirana kot evklidska razdalja oz. geometrična razdalja med položajem detekcije j in zadnjim veljavnim položajem trajektorije i :

$$C_{m,n} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,n} \end{bmatrix}, \quad c_{i,j} = d(t_i, d_j) = \sqrt{(x_{t_i} - x_{d_j})^2 + (y_{t_i} - y_{d_j})^2}$$

Rezultat Madžarske metode je permutiran vektor indeksov, ki predstavlja medsebojno prireditvev oz. povezavo detekcij in aktualnih trajektorij. Zatem je potrebno preveriti, če drži omejitev lokalnosti med položajem dodeljene detekcije in položajem izbrane trajektorije (parameter `cost_limit` v tabeli 6.1). Ta omejitev je preprosto pogoj, ki preveri, če se dodeljena detekcija nahaja dovolj blizu – v lokalni okolici – trajektorije. Madžarska metoda namreč poišče optimalno prirejanje glede na vrednosti v cenilni matriki, pri tem pa ne upošteva pogoja neposredne bližine. Prag omejitve določimo sorazmerno glede na predvideno hitrost tarč oz. pričakovano spremembo njihove pozicije (parameter `px_ch` v tabeli 6.1). V kolikor par detekcija-trajektorija ne ustreza omejitvi, se razveljavi indeks v prireditvenem vektorju, ki povezuje ta par. S tem povzročimo novo problematiko, saj ob razveljavitvi para določena trajektorija ostane brez detekcije, detekcija pa lahko pripada

novi trajektoriji. V nadaljevanju se zato dotaknemo tudi problema zastarelih in novih trajektorij.

Dodajanje novih trajektorij: Dodajanje poteka tako, da se najprej preveri, ali obstajajo detekcije, ki niso bile dodeljene obstoječim trajektorijam (zaradi omenjene razveljavitve ob neupoštevanem pogoju lokalnosti). V primeru, da take detekcije obstajajo, se jih dodeli povsem novim trajektorijam – pri tem velja, da postane trajektorija aktivna takrat, ko vsebuje vsaj eno detekcijo in ostaja aktivna vse dokler med detekcijami ne prihaja do večjih vrzeli. Pomen omejitve večjih vrzeli razložimo pri opisu brisanja zastarelih trajektorij. Tukaj lahko omenimo še, da lahko prihaja do težav v primeru lažnih detekcij, ker se s tem ustvarjajo tudi lažne trajektorije. Te so praviloma zelo kratke in jih je z dodatnimi omejitvami mogoče učinkovito odpravljati. Težave lahko predstavljajo tudi morebitni odboji oznak IR (v stiku z odbojnimi površinami scene), kar lahko poveča negotovost trajektorije, kateri pripada dotična detekcija.

Brisanje zastarelih trajektorij: Brisanje je osnovano na intuitivni ideji štetja manjkajočih detekcij. Za vsako trajektorijo tako implementiramo beleženje oz. evidenco aktivnosti. Ideja te evidence je dokaj enostavna. Predpostavimo namreč, da za trajektorije, ki obstajajo dlje časa, obstaja velika verjetnost, da bodo ostale aktivne dolgo časa tudi v prihodnosti. Velja tudi obratno. Torej za trajektorije, ki še ne obstajajo dolgo, obstaja določena verjetnost, da gre le za napačno asociacijo detekcij, zato omejimo njihov življenjski čas. Trenutna klasifikacija za trajektorijo T šteje manjkajoče detekcije T_{miss} ter njen življenjski čas T_{TTL} in ob izpolnitvi pogoja $T_{TTL} < T_{miss}$, se obravnavana trajektorija izbriše in ponastavi ter se postavi v vrsto, kjer je na voljo za oznake IR, ki se na novo pojavijo v sceni. Pogoj za izbris trajektorije je torej izpolnjen, če števec manjkajočih detekcij prekorači življenjsko dobo trajektorije (ki se poveča pri vsaki uspešno pridobljeni detekciji).

Ostale težave (lepljenje, prekrivanje, napake asociacije): Zaradi naivne logike brisanja trajektorij lahko pride do njihove fragmentacije. Ta pojav povzroči, da se zaradi negotovosti in manjkajočih detekcij več sledi ne združi v končno trajektorijo, čeprav ji te zagotovo pripadajo. Tukaj bi potrebovali izboljšavo metode brisanja ali pa dodaten mehanizem za lepljenje sledi, ki se niso uspele priključiti izbrani trajektoriji (kar pomeni izenačitev identitet obeh poti). Podobno se lahko pojavljajo težave ob prekrivanju ali

križanju trajektorij, kjer lahko prihaja do zamenjave identitet vpletenih objektov. Naša rešitev je uporaba sekundarnega sledilnika, ki preverja, ali je prišlo do nastanka fragmentacije. Če je fragmentacija prisotna se aktivira iskanje objekta na podlagi zadnje znane barvne informacije (ki je bila zajeta na koncu zaključene fragmentirane trajektorije). V kolikor sledenje v vidnem spektru pripelje do začetka druge fragmentirane trajektorije se trajektoriji združita in uskladita identiteto.

5.3.2 Sledilnik v vidnem spektru

Sledilnik v vidnem spektru preko sinhronizacije prejme delne rezultate iz primarnega sledilnika. Ker je njegova naloga povezovanje rezultatov iz spektra IR in preverjanje pravilnosti identitet, tukaj praktično ni potrebe po implementaciji avtonomnega detektorja. Kljub temu za zagotavljanje večje zmogljivosti sledilnika uporabimo zahtevnejši detektor na podlagi gradientnih značilnic, ki pa je na voljo opsijsko (v primeru, da imamo prosto grafično enoto). Velja poudariti, da se tukajšnje detekcije (v primerjavi z drugim sledilnikom) pojavljajo redkeje oz. so časovno zelo negotove (odvisno od zmogljivosti uporabljenega klasifikatorja). Namesto Kalmanovega filtra pri sledenju implementiramo metodo filtriranja z delci [45]. Zahtevnejša je tudi implementacija vizualnega modela, saj tukaj uporabimo utežene barvne histograme. Zaradi teh ključnih sprememb je sledilnik v vidnem spektru računsko in implementacijsko zahtevnejši. Ohranjanje identitet pa se tukaj nanaša na uporabo histogramov.

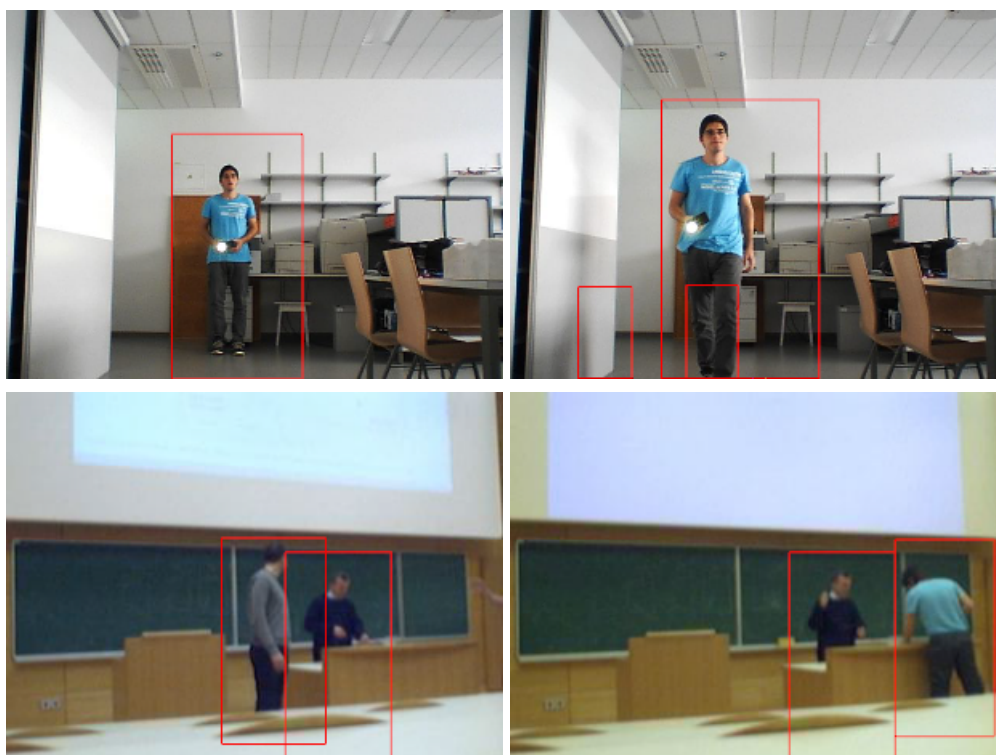
Detektor

Detektor sekundarnega sledilnika deluje na osnovi že omenjenih gradientnih deskriptorjev oz. značilnic HOG. Osnovna ideja teh deskriptorjev je ta, da izgled in obliko slikovne predloge opišejo z uporabo porazdeljenih lokalnih gradientov oz. robov. V praksi to pomeni, da se analizirana slika razdeli na majhne prostorske regije (celice). Zatem se za vsako izmed regij akumulira lokalni enodimenzionalni histogram, katerega stolpci (kategorije) predstavljajo različne smeri gradientov oz. robov. Če združimo histograme iz vseh celic, dobimo statistični opis gradientov celotne slike. Za večjo robustnost oz. invarianco osvetlitve (senčenje, lokalno različna osvetljenost) je priporočljiva predhodna normalizacija kontrasta v omenjenih regijah. Normalizacija vsake celice je zaradi upoštevanja večje okolice opravljena na večjih regijah oz. blokih (angl. *blocks*) [12].

V detekcijskem oknu algoritma se nato vzpostavi gosta mreža deskriptorjev HOG,

ki se med seboj prekrivajo. Na podlagi zajetih deskriptorjev se uporabi vnaprej naučen klasifikator SVM, ki presodi, ali množica slikovnih elementov v trenutnem detekcijskem oknu predstavlja osebo ali ne.

Glede na to, da je izvirni postopek za detekcijo t.i. metoda drsečega okna (angl. *sliding window*), hkrati pa izračun deskriptorjev ni trivialen, kmalu postane jasno, da izvirna sekvenčna implementacija tega algoritma ne deluje v realnem času. V naši nalogi smo zaradi realnočasovnih zahtev sistema uporabili obstoječo paralelno implementacijo¹ tega detektorja, imenovano **fastHOG** [46], ki najbolj zahtevno računsko delo detektorja preda grafični procesni enoti s pomočjo ogrodja Nvidia CUDA. Nekaj primerov detektiranih oseb je zbranih na sliki 5.6.



Slika 5.6 Primeri detektiranih oseb s fastHOG detektorjem (ena detekcija, lažne detekcije in več detekcij).

¹<https://github.com/ashwin/fasthog>

Sledilnik

Sledilnik sestavlja metoda filtriranja z delci [45], s katero se izvaja stohastično preiskovanje okolice v povezavi z vizualnim modelom v obliki barvnih histogramov. Poleg tega sledilnik vsebuje logiko za analizo in popravljanje rezultatov iz spektra IR in preverjanje identitet ter sinhronizacijo rezultatov s primarnim sledilnikom. Popravljeni rezultati pa predstavljajo tudi izhodne rezultate sistema.

Filter z delci: Za lokalno preiskovanje okolice trajektorij uporabimo verjetnostno metodo filtriranja z delci. Delec $p_i = \{x_i, y_i, v_{x_i}, v_{y_i}\}$ je entiteta, ki vsebuje informacijo o poziciji in hitrosti, poleg tega pa mu pripada tudi utež w_i , ki predstavlja njegovo pomembnost. Večja kot je utež delca, pomembnejši je delec, saj se nahaja bližje optimalni rešitvi. Naša implementacija uporablja konstantno število delcev (v nadaljevanju oznaka N , ki se nanaša na parameter `part_size` v tabeli 6.1), uteži pa so ob inicializaciji enakomerno utežene kot $w_i = \frac{1}{N}$. Proces filtriranja se izvede v 4 korakih: vzorčenje delcev, simulacija dinamičnega modela, izračun novih uteži ter izračun novega položaja tarče.

V prvem koraku delce prevzorčimo tako, da izvedemo izločanje glede na trenutne vrednosti njihovih uteži – pri tem se ohranijo delci z najboljšimi utežmi (elitizem). Uteži predhodno normaliziramo in sestavimo njihovo kumulativno distribucijo. Zatem z vzorčenjem iz enakomerne porazdelitve naključno izberemo N števil na intervalu $[0, 1]$. Za vsako naključno število zatem postavimo v kumulativno distribucijo ter si zapomnimo njegov indeks v tej distribuciji. Pridobljen seznam indeksov predstavlja izbiro elitnih delcev, ki preživijo fazo vzorčenja.

Korak simuliranja dinamičnega modela je dokaj preprost. Delce spustimo čez dinamični model, pri tem pa simulaciji dodamo (prištejemo) še negotovost, ki delce porazdeli po okolici preiskovanja. Dinamični model NCV je v matrični obliki definiran kot:

$$A_{NCV} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Celotni model simuliranja lahko zapišemo kot

$$p_{i,t} = A_{NCV} \cdot p_{i,t-1} + n_t; \quad n_t \sim \mathcal{N}(0, Q),$$

kjer je $p_{i,t}$ delec z indeksom i v času t , A_{NCV} je matrika dinamičnega modela, n_t pa je vektor z dodano negotovostjo, ki je definirana s kovariančno matriko Q .

Po simulaciji delcev je potrebno za vse delce p_i osvežiti vrednosti pripadajočih uteži w_i . Za vsak delec iz slike izrežemo slikovno predlogo, iz nje sestavimo barvni histogram P in ga primerjamo z obstoječim histogramom Q , ki pripada trajektoriji T . Primerjavo histogramov izvedemo z uporabo Hellingerjeve razdalje, ki je za dve diskretni distribuciji (v našem primeru histograma) P in Q definirana kot

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=0}^k (\sqrt{p_i} - \sqrt{q_i})^2},$$

kjer sta p_i in q_i vrednosti histogramov P in Q v stolpcu z indeksom i . Zaradi relacije z koeficientom *Bhattacharyya* (BC) lahko Hellingerjevo razdaljo zapišemo tudi kot:

$$H(P, Q) = \sqrt{1 - BC(P, Q)} = \sqrt{1 - \sum_{i=0}^k \sqrt{p_i \cdot q_i}}$$

Vrednost Hellingerjeve razdalje predstavlja novo vrednost uteži $w_i = H(P, Q)$ za izbrani delec p_i . V zadnji fazi izračunamo uteženo povprečje delcev oz. nov položaj objekta na trajektoriji T z izrazom

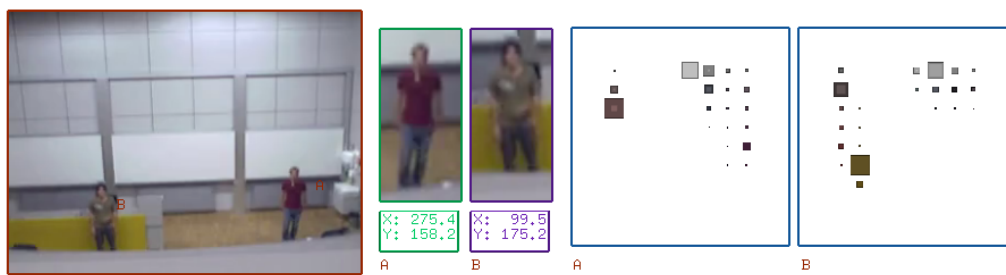
$$t_j = \frac{\sum_{i=0}^N p_i \cdot w_i}{\sum_{i=0}^N w_i},$$

kjer je t_j nova pozicija trajektorije T , na katero se nanašajo delci p_i in uteži w_i . Celoten algoritem se izvaja za vse aktualne oz. aktivne trajektorije, ki so fragmentirane in kot rečeno obravnava vsako trajektorijo posebej.

Vizualni model: Za vizualni model v vidnem delu uporabimo utežene barvne histograme. Utežitev je potrebna zaradi že znane predpostavke (iz [27]), da se detektiran objekt običajno nahaja v sredini izrezane slikovne predloge. To praktično pomeni, da moramo slikovnim elementom v sredini dodeliti večjo pomembnost, kot tistim elementom, ki se nahajajo na obrobju predloge. Na ta način ohranimo barvno informacijo objekta in nekoliko zatremo barvno informacijo iz ozadja. V nasprotju s pogosto uporabljenim Epanečnikovim jedrom (angl. *Epanechnikov kernel*), smo se zaradi hitrosti izračuna raje

odločili za uporabo trikotnega jedra (angl. *triangular kernel*). Vizualizacija dveh slikovnih predlog in pripadajočih histogramov je na sliki 5.7.

Poleg uteževanja uvedemo dodaten kriterij izločanja barv, za katere predvidevamo, da niso dovolj informativne za prepoznavanje oz. sledenje tarč. Pri tem predpostavimo, da je širina očitane pravokotnika tarče w_{bb} (parameter `w_bb` v tabeli 6.1) širša kot vitalni del same tarče w_t – ta pogoj lahko zapišemo kot $w_{bb} > w_t$. Če ta predpostavka velja, potem se neinformativne barve nahajajo tik ob levem in desnem robu slikovne zaplate, ki predstavlja osebo. Ob zunanjih vertikalnih robovih zaplate se zatem opravi predhodno vzorčenje barvnih odtenkov, ki se izvede pred vzorčenjem same zaplate. Polja v histogramu kamor spadajo predhodno vzorčeni odtenki postavimo na črno listo in jih v fazi vzorčenja celotne zaplate ne osvežujemo.



Slika 5.7 Primer vizualizacije dveh histogramov, generiranih na izrezanih predlogah.

Hkrati uvedemo še t.i. globalni histogram, ki naključno vzorči celotno sliko. Iz tega histograma lahko razberemo, katere barve se največkrat pojavijo. Barve, ki se pojavljajo velikokrat, običajno predstavljajo ozadje. Tudi v primeru, da predstavljajo del objektov, zaradi svojih številčnih pojavitev niso zelo informativne, zato jih lahko ponavadi zavržemo. Histogramu vsake detekcije oz. trajektorije zato odštejemo barve, ki so bile glede na globalni histogram izbrane kot slabo informativne. Delež teh barv določimo s parametrom `gh_thr` (tabela 6.1). Globalni histogram pa nam pove tudi, kakšna je globalna intenziteta osvetlitve na sceni. V primeru, da vrednost osvetlitve pade izven mejnega intervala (parametra `gh_high` in `gh_low` v tabeli 6.1), se sledenje na podlagi barvne informacije prekine in se aktivira šele, ko je nivo osvetljenosti ponovno na ustreznem intervalu.

Povezovanje trajektorij: Ker sledilnik upravlja z delnimi rezultati (fragmentiranimi trajektorijami) je potrebno logiko za upravljanje s temi trajektorijami ustrezno nadgra-

diti. Pri povezovanju trajektorij zato preverimo prostorsko in vizualno ujemanje novih detekcij in zaplat na zadnjih lokacijah obstoječih trajektorij. V primeru, da niti lokacijsko niti vizualno ne najdemo ujemanja med dvema trajektorijama, sklepamo, da nimata nič skupnega in jih med seboj ne povežemo.

Razlike nastanejo tudi pri brisanju trajektorij. Tukaj si pomagamo s primerjavo vizualnih uteži in uvedemo pogoj, s katerim glede na vrednosti uteži preverimo, ali je sledilnik zgrešil objekt ali ne. V primeru, da imajo uteži relativno majhne vrednosti (kar pomeni majhno ujemanje modela in preiskane okolice), potem trajektorijo brišemo in ponastavimo. Vse spremembe se zatem sinhronizirajo nazaj v primarni sledilnik, ki v naslednjem koraku sledenja uporabi popravljene podatke.

5.3.3 Sodelovanje sledilnikov

V tem predelu obstaja več možnosti za koordinacijo rezultatov iz obeh sledilnikov. Rezultati primarnega sledilnika (iz področja IR) so lahko uporabljeni kot nadomestek oz. dopolnitev detekcij v vidnem delu spektra. Tako predpostavimo, da se v okolici vsake detektirane regije v področju IR nahaja tudi pripadajoča oseba v vidnem spektru, ki nosi oznako IR (čeprav je mogoče detektor v vidnem spektru še ni prepoznal). Na ta način se zmanjša fragmentacija trajektorij v sledilniku vidnega spektra. V obratni smeri se uporabi barvno informacijo na ta način, da pri bližnji interakciji dveh ali večih oznak na relativno zanesljiv način preprečimo zamenjavo identitet (ob predpostavki, da lahko relativno zanesljivo razločimo osebe v vidnem spektru).

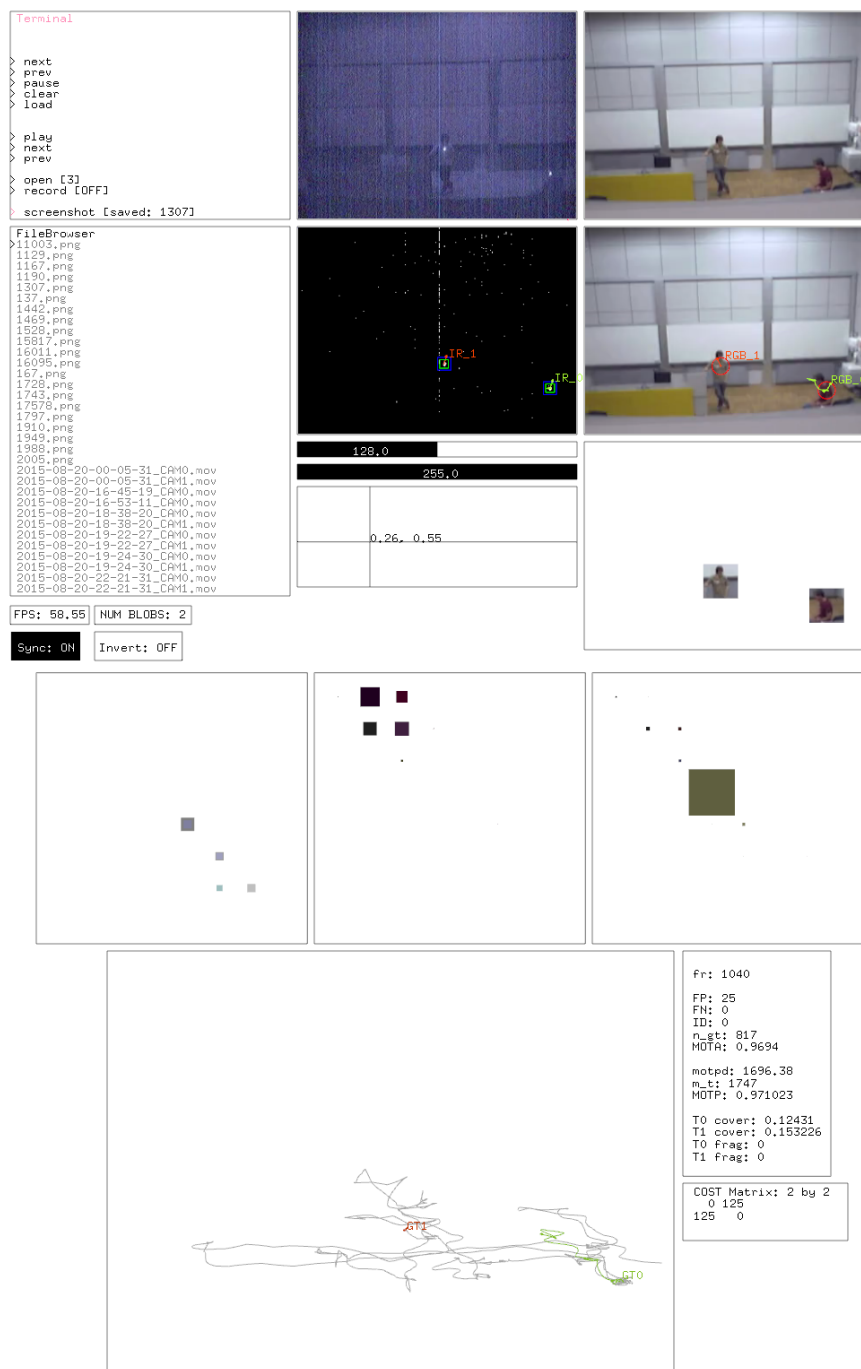
Ker se detekcije oseb v vidnem spektru relativno redko pojavljajo, je vodilno vlogo sledenja prevzel primarni sledilnik v spektru IR. Sledilnik v vidnem spektru pa je bil v dvo-stopenjski shemi (slika 5.4) usmerjen v dopolnitev in popravljanje rezultatov iz spektra IR. Primarni sledilnik najprej proizvede delne rezultate (sledi), ki so zaradi manjkajočih detekcij običajno fragmentirane – to pomeni, da jih ločujejo številne vrzeli. Ideja sodelovalne sheme je ta, da sekundarni (barvni) sledilnik skuša iz fragmentov rekonstruirati prave trajektorije in hkrati popraviti napake v identifikaciji, do katerih lahko pride med procesiranjem primarnega sledilnika. Na sliki 5.4 je sodelovanje predstavljeno v cikličnem sklopu, ki je sestavljen iz 4 zaporednih ukazov: 1. **sledi**, 2. **sinhroniziraj**, 3. **poveži** in 4. **sinhroniziraj nazaj**.

5.3.4 Uporabniški vmesnik

Uporabniški vmesnik (slika 5.8) zajema terminal (s predefiniranimi ukazi), brskalnik arhiviranih video sekvenc, okni za prikaz vhodnih slik v oba sledilnika, okni za prikaz rezultatov sledenja ter okna, ki ponujajo dodatne informacije (npr. prikaz cenilne matrike, prikaz trenutnih vrednosti evalvacijskih metrik, izris histogramov in anotacijske osnove ipd.). Terminal z majhnim naborom ukazov med drugim omogoča nalaganje sekvenc, koračno premikanje po izbrani sekvenci, nadzor predvajanja, inicializacijo priključenih kamer ter ponastavljanje trajektorij v trenutku.

5.3.5 Večnitnost sistema

Da bi zagotovili prilagodljivost in odzivno delovanje celotnega sistema, smo aplikacijo razdelili na več neodvisnih komponent, od katerih se nekatere izvajajo v ločenih procesorskih nitih. V glavni niti se izvajata oba sledilnika, detektor oznak IR in vse funkcionalnosti uporabniškega vmesnika. Sledilnik vidnega spektra lahko (opcijsko) zažene detektor oseb, ki se inicializira in izvaja v ločeni niti, procesiranje tega detektorja pa se – kot je bilo že omenjeno – izvede na grafični procesni enoti s pomočjo tehnologije CUDA. Za zajem video posnetkov iz posamezne kamere skrbi komponenta v ločeni niti, razlog pa je v tem, da je zajem slik lahko počasnejši od izvajanja glavne niti aplikacije. Skupno se torej izvajajo 3 niti (glavna nit in dve niti za zajem videa) ter opcijška nit (za detektor oseb v vidnem spektru).



Slika 5.8 Prikaz komponent uporabniškega vmesnika (terminal z ukazi, brskalnik arhiva posnetkov, vhodno in izhodno okno obeh sledilnikov, izpis cenilne matrike in evalvacijskih metrik ter prikaz histogramov in anotacijske osnove).

6 Vrednotenje sistema in rezultati

Klasični sledilniki običajno uporabljajo le informacijo iz vidnega spektra. V našem primeru uporabljamo informacijo iz dveh domen, zato klasičnih video sekvenc (npr. iz spletne baze¹, ki je bila zgrajena z namenom standardizacije vrednotenja) ni bilo mogoče uporabiti za vrednotenje in primerjavo obstoječih sledilnikov. Kljub temu pa (na lastnih scenarijih) izvedemo interno primerjavo (delnih) rezultatov sledilnika IR in končnih (združenih) rezultatov iz obeh sledilnikov.

Luo et al. [36] evalvacijske metrike klasificira glede na različne attribute (točnost, natančnost, popolnost in robustnost). Za merjenje točnosti omenja štetje zamenjanih identitet in kompleksnejšo metriko MOTA (predstavljena v [5]), ki kombinira število lažnih (napačnih) členov trajektorije (FP), zgrešene člene (FN) in število zamenjanih identitet (MME). Za merjenje natančnosti navede metriki TDE (iz [26]) in MOTP (iz [5]). Obe obravnavata natančnost sledenja glede na ujemanje (razdaljo ali presek) z anotacijsko osnovo. Metrike za popolnost prav tako klasificirajo trajektorije glede na ujemanje z anotacijsko osnovo, vendar pri tem preverjajo delež anotacije, ki ga dotična

¹https://motchallenge.net/data/2D_MOT_2015/

trajektorija pokrije (primer uporabe v [34]). Metrike robustnosti pa lahko služijo kot indikator, ki pove, kako uspešno si testiran sledilnik lahko opomore po kratkoročnem ali dolgoročnem prekrivanju (uporaba v [56]). Za izvedbo našega vrednotenja smo izbrali metrike MOTA, MOTP ter metrike za kategorizacijo popolnosti trajektorij.

Metrika MOTA je definirana kot

$$MOTA = 1 - \frac{\sum_t (fp_t + fn_t + mme_t)}{\sum_t g_t},$$

kjer je fp_t število napačno detektiranih členov vseh trajektorij v času t , fn_t je število zgrešenih členov vseh trajektorij v času t , mme_t pa predstavlja število dogodkov, pri katerih je prišlo do zamenjave identitet v času t . Člen g_t predstavlja število anotiranih trajektorij v času t . Zaloga vrednosti metrike MOTA zajema interval $(-\infty, 1]$. Tukaj vrednost 1 predstavlja najboljši možni rezultat (med sledenjem ni prišlo do napačnih dogodkov). Velja poudariti, da metrika ni navzdol omejena, kar pomeni, da je njena vrednost lahko negativna. V praksi se to lahko zgodi v primeru, kadar je skupno število neustreznih dogodkov (napak) bistveno večje od števila anotiranih trajektorij.

Metrika MOTP je definirana kot

$$MOTP = \frac{\sum_{i,t} \bar{d}(g_t^i, h_t^i)}{\sum_t c_t}; \quad \bar{d}(g_t^i, h_t^i) = 1 - \frac{\min(d(g_t^i, h_t^i), thr)}{thr},$$

kjer g_t^i predstavlja pozicijo zlatega standarda (GT) za objekt i v času t , h_t^i predstavlja pozicijo hipoteze (člen trajektorije) i v času t , operacija \bar{d} označuje mero podobnosti med anotacijo zlatega standarda in hipotezo, vrednost c_t je število anotiranih trajektorij v času t , operacija d je v našem primeru evklidska razdalja med dvema točkama, parameter thr pa označuje mejno vrednost razdalje ujemanja (v našem primeru je ta meja določena kot $thr = \frac{1}{2}w_{bb}$, kjer je w_{bb} širina očrtanega pravokotnika tarče – to je parameter sistema, ki je na voljo v tabeli 6.1). Zaloga vrednosti metrike MOTP zajema interval $[0, 1]$. Prav tako kot pri metriki MOTA, tudi tukaj vrednost 1 predstavlja najboljši možni rezultat (popolno ujemanje rezultatov z zlatim standardom).

Trajektorije posameznega scenarija klasificiramo še glede na metrike popolnosti, ki upoštevajo pokritost (popolnost) zlatega standarda (iz [56]). Kategorije so določene glede na delež pokritosti:

- večinsko pokrita (MT) – pokritost zlatega standarda nad 80%,

- večinsko izgubljena oz. zgrešena (ML) – pokritost zlatega standarda pod 20%,
- delno pokrita (PT) – pokritost zlatega standarda med 80% in 20%,
- ter z dodanim kriterijem fragmentiranosti (FM) – število dogodkov, ki spremenijo identiteto pokritega zlatega standarda.

Tako kot pri razvoju sistema, smo bili tudi pri izvedbi vrednotenja pozorni na naslednje probleme:

- odsotnost oznak IR pri prostem gibanju tarč,
- spreminjajoča se osvetlitev tekom sledenja (tema, delna zatemnitev, šibka osvetlitev in dobra osvetljenost prostora),
- interakcija dveh tarč pri omenjenih pogojih,
- pojav lažnih detekcij (zrcaljenje oznak IR v odbojnih površinah ozadja – tabla) in
- zmogljivost razločevanja tarč v vidnem spektru.

6.1 Izbira parametrov

Parametri so bili izbrani empirično, torej z opazovanjem delovanja sistema in njegove učinkovitosti pri različnih kombinacijah. Pri določanju vrednosti nekaterih parametrov smo upoštevali omejitve okolja v katerem so bili zajeti obravnavani scenariji (npr. omejitev največje možne hitrosti gibanja objektov, ki neposredno vpliva na vrednosti v kovariančni matriki negotovosti). Parametri sistema in njihove vrednosti, ki so bile uporabljene med evalvacijo so v tabeli 6.1. Poleg tega tabela vsebuje še informacijo o tem, v katerem sledilniku (IR ali vidnem – okrajšava RGB) posamezni parameter nastopa ter ponuja kratek opis parametra.

6.2 Scenariji eksperimentov

Scenarije, na katerih smo ovrednotili sistem, podajamo v nadaljevanju. V primerih nenačnne kalibracije kamere IR je bila napaka deloma kompenzirana z ustrezno translacijo objektov znotraj aplikacije, v izjemnih primerih (2. in 4. scenarij) pa tudi s predhodno rotacijo zajetega posnetka. S tem smo delne rezultate primarnega sledilnika poravnali z barvnimi značilnostmi iz vidnega spektra. Delne sledi in končne trajektorije so na slikah

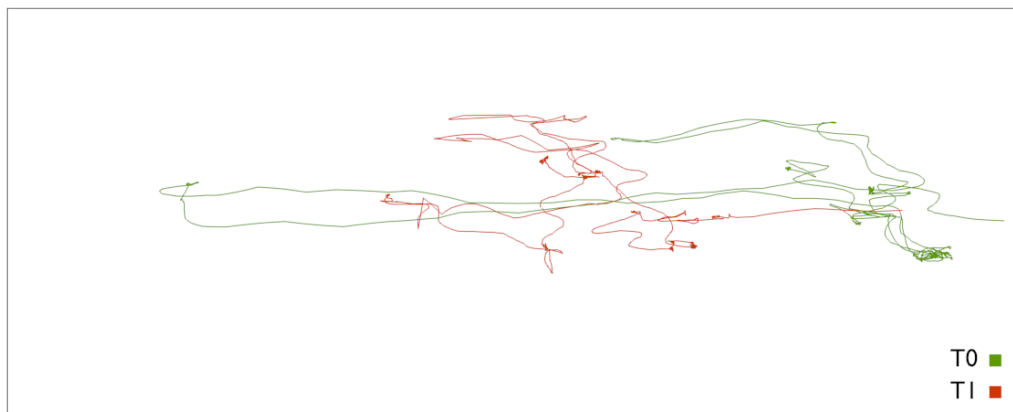
parameter	vrednost	sledilnik	opis	enota
thresh	128	IR	mejna vrednost IR detektorja	/
dt_ir	0.185	IR	časovni korak dinamičnega modela	/
dt_rgb	1.0	RGB	časovni korak dinamičnega modela	/
px_ch	5	oba	pričakovana sprememba pozicije	[px]
cost_limit	15	IR	zgornja meja razdalje asociacije	[px]
w_bb	40	RGB	širina pravokotnika tarče	[px]
h_bb	40	RGB	širina pravokotnika tarče	[px]
link_limit	30	RGB	zgornja meja razdalje povezovanja	[px]
redu_limit	10	RGB	zgornja meja razdalje odstranjevanja	[px]
hist_limit	85	RGB	spodnja meja ujemanja histogramov	[%]
hist_size	8	RGB	število razdelkov (stolpcev) histograma	/
gh_thr	80	RGB	delež barv ozadja v globalnem histogramu	[%]
gh_high	200	RGB	zgornja meja osvetljenosti	/
gh_low	120	RGB	spodnja meja osvetljenosti	/
part_size	50	RGB	število delcev	/

Tabela 6.1 Vrednosti parametrov sistema, ki so bile izbrane za evalvacijo. Časovni korak dinamičnega modela v sledilniku IR kompenzira problematično hitrost osveževanja kamere IR.

označene z različnimi kontrastnimi barvami ter nosijo ustrezno oznako T_{id} , kjer oznaka id predstavlja identifikacijsko številko trajektorije T .

Referenčne poti objektov so bile rekonstruirane s kombiniranjem rezultatov iz spektra IR in ročnih anotacij preostalega dela video sekvenc. Na ta način smo pridobili zlati standard, ki predstavlja gibanje nastopajočih objektov v realnosti. Vizualizacije zlatega standarda vseh scenarijev (skupaj z referenčnim okvirjem prizorišča) so zbrane v slikah [6.1](#), [6.2](#), [6.3](#) in [6.4](#).

Za razvoj sistema smo uporabljali ločene scenarije, kjer so bile testne sekvence sicer krajše, vendar so naslovile enake probleme kot scenariji, ki so bili obravnavani v končnem vrednotenju (sprememba osvetlitve, interakcija udeležencev, odhajanje iz prizorišča ipd.).



Slika 6.1 Vizualizacija zlatega standarda za 1. scenarij.

6.2.1 Scenarij 1: Gibanje dveh oseb pri konstantni osvetljenosti

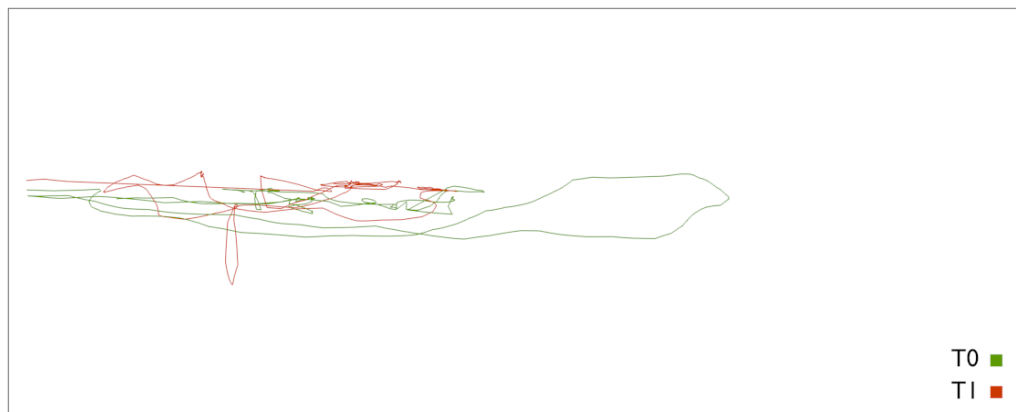
V tem scenariju je predstavljeno prosto gibanje dveh oseb, kjer prihaja do križanja poti udeležencev. Spremembe v osvetlitvi prizorišča ni, zaradi odbojev markerja od table v ozadju pa se v področju IR pojavijo lažne detekcije. Te aktivirajo kratko trajajoče lažne trajektorije, ki pa ne povzročajo motenj na obstoječih (veljavnih) trajektorijah.

6.2.2 Scenarij 2: Interakcija dveh oseb pri konstantni osvetljenosti in gibanje po zatemnjenem prizorišču

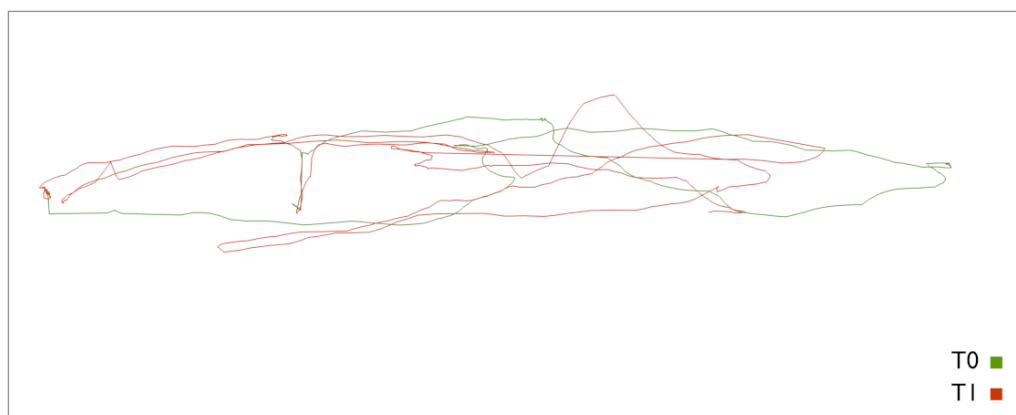
Drugi scenarij uvede dve novi situaciji. V prvi polovici scenarija se prikaže bližnja interakcija dveh oseb, ki se odvija pri konstantni osvetljenosti. V drugi polovici sledi zatemnitev prizorišča in gibanje oseb v delni temi. Vidimo, da se pri sledenju pojavi precej težav, dodatne probleme pa (med interakcijo oseb) predstavlja tudi ojačan (direktno usmerjen) žarek infrardeče svetlobe, ki se kaže kot skupina koncentriranih regij, kar povzroči nastanek gruče lažnih detekcij. Ta gruča povzroči fragmentacijo identitet obeh objektov.

6.2.3 Scenarij 3: Izmenično gibanje in mirovanje dveh oseb z vmesnimi zatemnitvami

Tretji scenarij v prvi polovici trajanja prikaže gibanje ene osebe pri večkratni spremembi osvetlitve (medtem druga oseba čaka). Zatem se v drugi polovici vlogi gibanja zamenjata. Scenarij ima namen preveriti, ali sledilnik zmore ohranjati identiteto udeležencev med zamenjavo mirujoče in gibajoče se osebe. Izjemne spremembe osvetljenosti povzročijo,



Slika 6.2 Vizualizacija zlatega standarda za 2. scenarij.

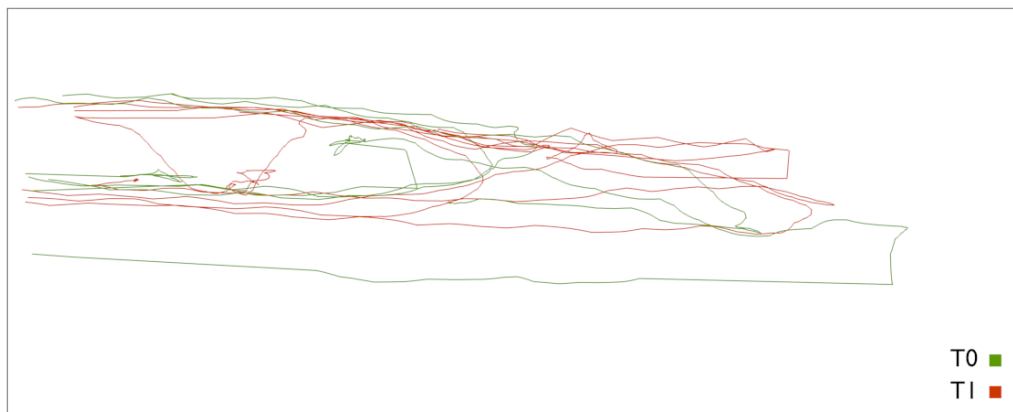


Slika 6.3 Vizualizacija zlatega standarda za 3. scenarij.

da je sledilnik v vidnem spektru nekaj časa popolnoma onemogočen. Prisoten šum v področju IR pa generira lažne detekcije, ki v prvi tretjini scenarija aktivirajo kaskado lažnih trajektorij. Scenarij pokaže tudi slabost kamere IR, ki reagira na sunkovito spremembo osvetlitve in tako povzroči množico lažnih detekcij.

6.2.4 Scenarij 4: Pojavljanje in izginjanje dveh oseb z uvodno zatemnitvijo in vmesnimi spremembami osvetljenosti

V zadnjem scenariju osebi na začetku krožita okoli poligona pri nizki globalni osvetljenosti ter se v nadaljevanju prosto gibljeta. Dodatno težavnost scenarija predstavlja dejstvo, da tarči nekajkrat popolnoma zapustita prizorišče in se zatem vrneta nazaj. Osvetljenost prizorišča se iz začetne nizke vrednosti nekajkrat v kratkem času povzpne in zatem po-



Slika 6.4 Vizualizacija zlatega standarda za 4. scenarij.

novno pade na izvorno raven. Zaradi zatemnjenega začetka sledilnik vidnega spektra ne nudi opore, dokler se osvetljenost ne popravi. To povzroči večje število neodpravljenih napak v predelu sledenja spektra IR. Ta scenarij je zaradi svojih pogojev (izginjanja tarč in začetne nizke osvetlitve) zelo težaven, kar se kaže tudi v rezultatih.

6.3 Rezultati

Kvaliteto sledenja smo določili glede na ujemanje dobljenih rezultatov in zlatega standarda. Ker je sledenje v vidnem delu spektra stohastično, smo postopke evalvacije izvedli večkrat (10 ponovitev) in upoštevali povprečne vrednosti ter standardni odklon. Rezultati sledenja v spektru IR so zbrani v tabelah 6.2 in 6.3. Rezultati sledenja v spektru IR z dodatno pomočjo sledilnika v vidnem spektru so zbrani v tabelah 6.4 in 6.5.

Delnih rezultatov, ki bi zajemali le delovanje vidnega sledilnika ni na voljo, ker je ta implementiran tako, da na vhodu pričakuje delne rezultate iz spektra IR. Zaradi tega razloga vidnega sledilnika ni mogoče pognati samostojno. Intuitivno pa vemo, da bi sledilnik v vidnem delu odpovedal med vsemi zatemnjenimi deli scenarijev, zaradi česar so njegovi rezultati zagotovo slabši od delnih rezultatov sledenja iz spektra IR.

V 1. scenariju metriki točnosti (MOTA) in natančnosti (MOTP) pri obeh načinih sledenja zavzameta visoko vrednost, zelo visoka je tudi popolnost (pokritost) zlatega standarda pri obeh trajektorijah, fragmentacije je malo, sledenje v vidnem spektru pa jo še dodatno zmanjša (tabela 6.2 in 6.4). V 2. scenariju sta zaradi gruč lažnih detekcij oceni MOTA in MOTP nekoliko nižji. Pokritost zlatega standarda je relativno uspešna,

fragmentacija pa je na obeh trajektorijah višja kot v 1. scenariju. Čeprav sta metriki MOTA in MOTP pri vključenem sledenju vidnega dela nekoliko nižji, se prednost vidnega sledenja kaže v občutnem zmanjšanju fragmentacije (tabela 6.4). V 3. scenariju prihaja do sunkovite spremembe osvetlitve, kar se pokaže tudi v izredno nizki vrednosti metrike MOTA v obeh različicah sledenja (tabela 6.2 in 6.4), vrednost metrike MOTP je sicer nekoliko višja, vendar nižja kot v 2. scenariju. Fragmentacija je manjša v trajektoriji T_0 in izrazita v T_1 . Tudi tukaj vključeno sledenje v vidnem spektru dodatno zmanjša prisotno fragmentacijo. V 4. scenariju dodatne probleme povzroča prekinitev trajektorij zaradi izginjanja oseb, zaradi česar je prisotna visoka fragmentacija. Poleg tega sta vrednosti MOTA in MOTP za oba načina sledenja relativno nizki (čeprav boljši kot v 3. scenariju), šibka pa je tudi mera popolnosti pri obeh trajektorijah (tabela 6.2 in 6.4).

Kvalitativne ugotovitve vrednotenja pa so naslednje:

- sledenje v spektru IR deluje zadovoljivo, kadar so detekcije iz spektra IR pravilne,
- sledenje v vidnem spektru deluje zadovoljivo, kadar so tarče vizualno dovolj različne,
- v splošnem so združeni rezultati boljši kot rezultati sledenja IR,
- sledenje v vidnem spektru občutno zmanjša fragmentacijo trajektorij,
- hkratna odsotnost osvetlitve in detekcij v spektru IR povzroči, da sledenje odpove,
- detektor v spektru IR lahko zaradi šuma, sunkovite intenzitetne spremembe (3. in 4. scenarij) ali prisotnih odbojnih površin generira lažne detekcije,
- do zamenjave identitet pride v primeru medsebojne interakcije, kadar v vidnem spektru nimamo dovolj vizualnih informacij o vpletenih tarčah (zaradi nizkega faktorja osvetlitve ali popolnega prekrivanja),
- dolgotrajna odsotnost detekcij iz spektra IR prepusti preiskovanje vidnemu delu, ki lahko (zaradi spremenljive osvetlitve) tarčo izgubi.

Zaželjena bi bila še primerjava s sorodnimi pristopi, vendar tega zaradi manjkajočih vhodnih podatkov (kanala IR) in različnega namena uporabe (npr. zunaj ali v prometu) ne bi bilo mogoče izvesti na enaki bazi posnetkov. Vseeno so zgolj za informativnost v tabeli 6.6 priloženi rezultati nekaterih sorodnih pristopov v vidnem spektru, ki so bili testirani na preizkusu 2D MOT 2015².

²https://motchallenge.net/results/2D_MOT_2015/

Scenarij	\overline{MOTA} [%] ↑	\overline{MOTP} [%] ↑	$\overline{popolnost}$ [%] ↑	\overline{FM} [#] ↓
1.	82.96	92.72	T_0 : 83.46, T_1 : 87.12	T_0 : 3.50, T_1 : 5.30
2.	62.10	82.18	T_0 : 55.44, T_1 : 70.32	T_0 : 12.90, T_1 : 6.30
3.	12.99	63.94	T_0 : 29.54, T_1 : 70.59	T_0 : 5.00, T_1 : 15.50
4.	21.24	70.85	T_0 : 44.24, T_1 : 46.55	T_0 : 16.40, T_1 : 6.00

Tabela 6.2 Povprečne vrednosti evalvacijskih metrik MOTA, MOTP, popolnosti in fragmentiranosti za rezultate iz spektra IR. Oznaka ↑ pomeni, da so višje vrednosti metrike boljše od nižjih. Pri oznaki ↓ velja obratno.

Scenarij	σ MOTA	σ MOTP	σ popolnost	σ FM
1.	0.34	0.19	T_0 : 0.17, T_1 : 0.09	T_0 : 0.81, T_1 : 2.97
2.	0.70	0.20	T_0 : 0.17, T_1 : 0.35	T_0 : 1.58, T_1 : 1.79
3.	2.22	1.41	T_0 : 0.21, T_1 : 1.20	T_0 : 0.00, T_1 : 1.50
4.	1.38	0.24	T_0 : 1.07, T_1 : 0.65	T_0 : 3.95, T_1 : 1.26

Tabela 6.3 Standardni odkloni evalvacijskih metrik MOTA, MOTP, popolnosti in fragmentiranosti za rezultate iz spektra IR.

Scenarij	\overline{MOTA} [%] ↑	\overline{MOTP} [%] ↑	$\overline{popolnost}$ [%] ↑	\overline{FM} [#] ↓
1.	84.10	93.00	T_0 : 83.43, T_1 : 85.73	T_0 : 0.80, T_1 : 0.10
2.	55.93	78.01	T_0 : 55.36, T_1 : 63.46	T_0 : 1.90, T_1 : 3.70
3.	13.17	65.25	T_0 : 29.33, T_1 : 70.68	T_0 : 2.00, T_1 : 6.00
4.	26.61	74.24	T_0 : 49.89, T_1 : 51.09	T_0 : 4.80, T_1 : 3.80

Tabela 6.4 Povprečne vrednosti evalvacijskih metrik MOTA, MOTP, popolnosti in fragmentiranosti za združene rezultate iz obeh sledilnikov. Oznaka ↑ pomeni, da so višje vrednosti metrike boljše od nižjih. Pri oznaki ↓ velja obratno.

Scenarij	σ MOTA	σ MOTP	σ popolnost	σ FM
1.	2.48	0.85	T_0 : 3.89, T_1 : 2.65	T_0 : 0.98, T_1 : 0.30
2.	1.04	0.46	T_0 : 1.18, T_1 : 0.71	T_0 : 0.83, T_1 : 0.46
3.	0.56	0.68	T_0 : 0.54, T_1 : 1.04	T_0 : 0.00, T_1 : 1.61
4.	2.41	1.29	T_0 : 2.74, T_1 : 2.44	T_0 : 1.47, T_1 : 1.54

Tabela 6.5 Standardni odkloni evalvacijskih metrik MOTA, MOTP, popolnosti in fragmentiranosti za združene rezultate iz obeh sledilnikov.

Tukaj je potrebno poudariti, da je le en sledilnik v tabeli (sledilnik SORT) javno dostopen, pri vseh ostalih izvorna koda ni na voljo. Če bi želeli sledilnik SORT ovrednotiti na vidnem delu naših scenarijev, bi bilo potrebno vanj dodati programsko kodo za nalaganje zlatega standarda ter izračun vseh uporabljenih metrik. Sama implementacija sledilnika SORT ne vsebuje te logike, saj preizkus 2D MOT 2015 ponuja neodvisno ogrodje za vrednotenje. Kljub temu iz rezultatov v tabeli 6.6 lahko razberemo, da je sledilnik SORT trenutno eden izmed boljših sprotnih sledilnikov v vidnem delu, ki imajo hitrost izvajanja primerno za izvajanje v realnem času (kar je primerljivo z našim sistemom).

Sledilnik	MOTA [%] ↑	MOTP [%] ↑
SORT[6]	33.4	72.1
lkct_mot	32.6	72.7
MTRM	27.7	70.7
STP_MAP	25.7	70.8
OMT_DFH	21.2	69.9
RNN_LSTM	19.0	71.0
TSDA_OAL	18.6	69.7
HKTC	17.8	70.6
GSCR[14]	15.8	69.4
LDCT[55]	4.7	71.7

Tabela 6.6 Vrednosti evalvacijskih metrik MOTA in MOTP za sorodne sekvenčne sledilnike s primerljivo hitrostjo izvajanja, ki so bili testirani na preizkusu 2D MOT 2015. Oznaka ↑ pomeni, da so višje vrednosti metrike boljše od nižjih. Zaradi prisotnega anonimnega sodelovanja ali pomanjkljivih podatkov na preizkusu nekatere reference niso na voljo.

V prilogah A so na voljo še slike A.1, A.3, A.5, A.7, kjer so prikazani rezultati sledenja v spektru IR ter slike A.2, A.4, A.6, A.8, kjer je povzet potek sledenja v vidnem delu. Oba prikaza sta v obliki slikovnega zaporedja, ki zajema ključno dinamiko posameznega scenarija.

7 Zaključek

S tem poglavjem zaključujemo raziskovalno nalogo. V nadaljevanju navedemo glavne prispevke, ki jih prinaša to raziskovalno delo. Za konec pa predlagamo še nekaj smernic, s katerimi bi lahko nadgradili oz. izboljšali trenutne rezultate.

7.1 Glavni prispevki

Implementirali smo dva sledilna algoritma, ki sta namenjena sledenju objektov v obliki intenzitetnih regij ter barvnih predlog. Pri tem so bile uporabljene dobro poznane in dokumentirane metode na področju računalniškega vida.

Kot glavni prispevek bi izpostavili združitev delovanja dveh ločenih sledilnikov, ki imata možnost medsebojnega dopolnjevanja. Čeprav je trenutna implementacija medsebojnega sodelovanja dokaj osnovna, se je na podlagi praktičnih testov v dvorani izkazalo, da pomoč sledilnika v spektru IR izrazito izboljša točnost sledenja (tukaj velja omeniti, da smo zaradi tega primorani uporabiti dodatne svetlobne oznake, ki področje možne uporabe nekoliko omejujejo). Poleg tega na nivoju posameznega sledilnika pokažemo, kako združiti metode za sledenje enega objekta in metode podatkovne asociacije tako, da se

omogoči sledenje večih objektov hkrati.

Preostali manjši prispevki pa zajemajo idejne rešitve, ki smo jih zaradi upoštevanja vnaprej določenih predpostavk vključili v proces sledenja (npr. uvedba kriterija življenjske dobe trajektorije, uvedba globalnega histograma za določanje globalne osvetljenosti ter implementacijsko hiter način izločanja barv ozadja iz histogramov tarč z upoštevanjem informacije iz globalnega histograma).

Razvoj sistema je poleg implementacije samih sledilnih algoritmov zajemal tudi skrbno arhitekturno načrtovanje celotnega systemskega ogrodja in integracijo preostalih zahtevanih komponent (interaktiven uporabniški vmesnik in vizualizacija, komunikacija z različnimi modeli kamer, vgradnja obstoječega detektorja v vidnem delu ter podpora za večnitnost).

7.2 Možnosti nadaljnjega raziskovanja

Opisana implementacija sistema vsekakor odpira nove možnosti za izboljšave. Na izboljšanje rezultatov lahko praktično vpliva vsaka posamezna komponenta, ki je integrirana v sistem. Tako bi lahko npr. izboljšali delovanje detektorja regij v področju IR na dva načina. Prvi način predstavlja izboljšavo same implementacije detektorja (filtriranje napačnih regij, učenje izgleda pravih regij, zatiranje sunkovitih sprememb v velikosti zaznanih svetlobnih oznak ipd.), drugi način pa lahko zajema nakup kamer višjega razreda, s katerimi bi lahko izboljšali doseg ter občutljivost detektorja (kar je poglavito za uporabo sistema v večjih prostorih). Sem lahko spada tudi sprememba uporabe svetlobnih oznak, predvsem v smislu števila uporabljenih svetlobnih virov in načina njihovega nošenja.

Detektor v vidnem delu spektra bi lahko izboljšali z zamenjavo vgrajenega klasifikatorja SVM. Trenutno znanje klasifikatorja namreč izvira iz splošne slikovne baze pešcev. Prepoznavanje oseb je zato na splošno povsem zadovoljivo, vendar predvidevamo, da bi lahko ta rezultat izboljšali z dopolnjenim klasifikatorjem, ki bi bil naučen iz množice specifičnih učnih primerov. Ti primeri bi bili ustrezno pridobljeni iz scenarijev, ki so bližje področju uporabe našega sistema.

Oba sledilnika trenutno objekte v fazi sledenja obravnavata ločeno, zatem pa združujeta bližnje detekcije oz. podobne delne trajektorije. Pristop bi bilo zanimivo zamenjati z metodami, ki stanja objektov obravnavajo skupno ter primerjati rezultate sledenja iz

obeh pristopov. Posebno pazljivost bi lahko namenili tudi bolj premišljeni izbiri parametrov, npr. z uporabo formalnejših metod. Trenutno namreč večino parametrov določa t.i. pravilo palca oz. povsem praktični empirični preizkusi z opazovanjem delovanja sistema.

Pri sledilniku v vidnem spektru bi bilo potrebno več pozornosti posvetiti tudi vizualnemu modelu objektov. Pri tem bi lahko preizkusili različna jedra za uteževanje histogramov, preverili učinkovitost uporabe različnih barvnih prostorov ter izboljšali modeliranje barvnih odtenkov ozadja. Konkretna izboljšava pa bi lahko bil tudi sestavljen vizualni model, kjer bi upoštevali tudi npr. gradientne značilke.

Velja omeniti še področje sodelovanja obeh sledilnikov. Tukaj bi bila zaželjena implementacija naprednejših metod za primerjanje in koordiniranje rezultatov obeh sledilnikov. Metode v tem sklopu bi predvsem usklajevale stanje aktivnih trajektorij v obeh sledilnikih ter inteligentno odpravljale morebitne napake, ki jih iz vidika posameznega sledilnika ni mogoče zaznati oz. odpraviti.

LITERATURA

- [1] S. Ali, M. Shah, Floor fields for tracking in high density crowd scenes, in: Proceedings of the 10th European Conference on Computer Vision: Part II, ECCV '08, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 1–14.
url: http://dx.doi.org/10.1007/978-3-540-88688-4_1
- [2] A. Andriyenko, K. Schindler, S. Roth, Discrete-continuous optimization for multi-target tracking, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 1926–1933.
- [3] A. Andriyenko, K. Schindler, Multi-target tracking by continuous energy minimization, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011, pp. 1265–1272.
- [4] J. Berclaz, F. Fleuret, E. Turetken, P. Fua, Multiple object tracking using k-shortest paths optimization, Pattern Analysis and Machine Intelligence, IEEE Transactions on 33 (9) (2011) 1806–1819.
- [5] K. Bernardin, R. Stiefelhagen, Evaluating multiple object tracking performance: The clear mot metrics, J. Image Video Process. 2008 (2008) 1:1–1:10.
url: <http://dx.doi.org/10.1155/2008/246309>
- [6] A. Bewley, Z. Ge, L. Ott, F. Ramos, B. Upcroft, Simple online and realtime tracking, arXiv:1602.00763.
url: <http://arxiv.org/abs/1602.00763>
- [7] M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, L. Van Gool, Robust tracking-by-detection using a detector confidence particle filter, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 1515–1522.

- [8] W. Brendel, M. Amer, S. Todorovic, Multiobject tracking as maximum weight independent set, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011, pp. 1273–1280.
- [9] G. J. Brostow, R. Cipolla, Unsupervised bayesian detection of independent motion in crowds, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 1, 2006, pp. 594–601.
- [10] A. A. Butt, R. T. Collins, Multi-target tracking by lagrangian relaxation to min-cost network flow, in: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, 2013, pp. 1846–1853.
- [11] W. Choi, S. Savarese, Multiple target tracking in world coordinate with single, minimally calibrated camera, in: Proceedings of the 11th European Conference on Computer Vision: Part IV, ECCV'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 553–567.
url: <http://dl.acm.org/citation.cfm?id=1888089.1888132>
- [12] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1, 2005, pp. 886–893 vol. 1.
url: <http://dx.doi.org/10.1109/CVPR.2005.177>
- [13] A. Ess, B. Leibe, K. Schindler, L. Van Gool, Robust multiperson tracking from a mobile platform, Pattern Analysis and Machine Intelligence, IEEE Transactions on 31 (10) (2009) 1831–1846.
- [14] L. Fagot-Bouquet, R. Audigier, Y. Dhome, F. Lerasle, Online multi-person tracking based on global sparse collaborative representations, in: Image Processing (ICIP), 2015 IEEE International Conference on, 2015, pp. 2414–2418.
- [15] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, Pattern Analysis and Machine Intelligence, IEEE Transactions on 32 (9) (2010) 1627–1645.
- [16] F. Fleuret, J. Berclaz, R. Lengagne, P. Fua, Multi-camera people tracking with a probabilistic occupancy map, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (2) (2008) 267–282.

- [17] D. M. Gavrilă, S. Munder, Multi-cue pedestrian detection and tracking from a moving vehicle, *Int. J. Comput. Vision* 73 (1) (2007) 41–59.
url: <http://dx.doi.org/10.1007/s11263-006-9038-7>
- [18] J. Giebel, D. M. Gavrilă, C. Schnorr, A bayesian framework for multi-cue 3d object tracking, in: *In Proceedings of European Conference on Computer Vision*, Springer-Verlag, 2004, pp. 241–252.
- [19] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, *Phys. Rev. E* 51 (1995) 4282–4286.
- [20] J. F. Henriques, R. Caseiro, J. Batista, Globally optimal solution to multi-object tracking with merged measurements., in: D. N. Metaxas, L. Quan, A. Sanfeliu, L. J. V. Gool (Eds.), *ICCV*, IEEE Computer Society, 2011, pp. 2470–2477.
url: <http://dx.doi.org/10.1109/ICCV.2011.6126532>
- [21] W. Hu, X. Li, W. Luo, X. Zhang, S. Maybank, Z. Zhang, Single and multiple object tracking using log-euclidean riemannian subspace and block-division appearance model, *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* 34 (12) (2012) 2420–2440.
- [22] M. Hu, S. Ali, M. Shah, Detecting global motion patterns in complex videos, in: *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 2008, pp. 1–5.
- [23] H. Izadinia, I. Saleemi, W. Li, M. Shah, (mp)_i²/_i^t: Multiple people multiple parts tracker., in: A. W. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid (Eds.), *ECCV* (6), Vol. 7577 of *Lecture Notes in Computer Science*, Springer, 2012, pp. 100–114.
url: http://dx.doi.org/10.1007/978-3-642-33783-3_8
- [24] C. Kim, F. Li, A. Ciptadi, J. M. Rehg, Multiple hypothesis tracking revisited., in: *ICCV*, IEEE, 2015, pp. 4696–4704.
- [25] L. Kratz, K. Nishino, Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes, *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* 34 (5) (2012) 987–1002.

- [26] L. Kratz, K. Nishino, Tracking with local spatio-temporal motion patterns in extremely crowded scenes, in: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, 2010, pp. 693–700.
- [27] M. Kristan, J. Perš, M. Perše, S. Kovačič, Closed-world tracking of multiple interacting targets for indoor-sports applications, Computer Vision and Image Understanding 113 (5) (2009) 598–611.
- [28] H. W. Kuhn, The hungarian method for the assignment problem, Naval Research Logistics Quarterly 2 (1955) 83–97.
- [29] K. Kumar, C. De Vleeschouwer, Discriminative label propagation for multi-object tracking with sporadic appearance features, in: Computer Vision (ICCV), 2013 IEEE International Conference on, 2013, pp. 2000–2007.
- [30] C.-H. Kuo, C. Huang, R. Nevatia, Multi-target tracking by on-line learned discriminative appearance models., in: CVPR, IEEE Computer Society, 2010, pp. 685–692. url: <http://doi.ieeecomputersociety.org/10.1109/CVPR.2010.5540148>
- [31] C.-H. Kuo, R. Nevatia, How does person identity recognition help multi-person tracking?, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011, pp. 1217–1224.
- [32] S. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, in: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, CVPR '06, IEEE Computer Society, Washington, DC, USA, 2006, pp. 2169–2178. url: <http://dx.doi.org/10.1109/CVPR.2006.68>
- [33] B. Leibe, K. Schindler, N. Cornelis, L. Van Gool, Coupled object detection and tracking from static cameras and moving vehicles, Pattern Analysis and Machine Intelligence, IEEE Transactions on 30 (10) (2008) 1683–1698.
- [34] Y. Li, C. Huang, R. Nevatia, Learning to associate: Hybridboosted multi-target tracker for crowded scene, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 2953–2960.

- [35] W.-L. Lu, J.-A. Ting, J. J. Little, K. P. Murphy, Learning to track and identify players from broadcast sports videos, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (7) (2013) 1704–1716.
url: <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2012.242>
- [36] W. Luo, X. Zhao, T.-K. Kim, Multiple object tracking: A review, *CoRR* abs/1409.7618.
url: <http://arxiv.org/abs/1409.7618>
- [37] W. Luo, B. Stenger, X. Zhao, T.-K. Kim, Automatic topic discovery for multi-object tracking, in: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI'15*, AAAI Press, 2015, pp. 3820–3826.
url: <http://dl.acm.org/citation.cfm?id=2888116.2888246>
- [38] A. Milan, S. Roth, K. Schindler, Continuous energy minimization for multitarget tracking, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 36 (1) (2014) 58–72.
- [39] A. Milan, K. Schindler, S. Roth, Detection- and trajectory-level exclusion in multiple object tracking, in: *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 3682–3689.
- [40] D. Mitzel, E. Horbert, A. Ess, B. Leibe, Multi-person tracking with sparse detection and continuous segmentation., in: K. Daniilidis, P. Maragos, N. Paragios (Eds.), *ECCV* (1), Vol. 6311 of *Lecture Notes in Computer Science*, Springer, 2010, pp. 397–410.
url: http://dx.doi.org/10.1007/978-3-642-15549-9_29
- [41] P. Mordohai, G. Medioni, Dimensionality estimation, manifold learning and function approximation using tensor voting, *J. Mach. Learn. Res.* 11 (2010) 411–450.
url: <http://dl.acm.org/citation.cfm?id=1756006.1756018>
- [42] S. Pellegrini, A. Ess, L. Van Gool, Improving data association by joint modeling of pedestrian trajectories and groupings, in: *Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV'10*, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 452–465.
url: <http://dl.acm.org/citation.cfm?id=1886063.1886098>

- [43] S. Pellegrini, A. Ess, K. Schindler, L. Van Gool, You'll never walk alone: Modeling social behavior for multi-target tracking, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 261–268.
- [44] F. Porikli, O. Tuzel, P. Meer, Covariance tracking using model update based on lie algebra, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 1, 2006, pp. 728–735.
- [45] S. Prince, Computer Vision: Models Learning and Inference, Cambridge University Press, 2012.
- [46] V. Prisacariu, I. Reid, fasthog - a real-time gpu implementation of hog, Tech. Rep. 2310/09, Department of Engineering Science, Oxford University.
- [47] P. Pérez, C. Hue, J. Vermaak, M. Gangnet, Color-based probabilistic tracking, in: Proceedings of the 7th European Conference on Computer Vision-Part I, ECCV '02, Springer-Verlag, London, UK, UK, 2002, pp. 661–675.
url: <http://dl.acm.org/citation.cfm?id=645315.649166>
- [48] Z. Qin, S. R. Christian, Improving multi-target tracking via social grouping, in: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), CVPR '12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 1972–1978.
url: <http://dl.acm.org/citation.cfm?id=2354409.2354969>
- [49] V. Reilly, H. Idrees, M. Shah, Detection and tracking of large number of targets in wide area surveillance, in: Proceedings of the 11th European Conference on Computer Vision Conference on Computer Vision: Part III, ECCV'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 186–199.
url: <http://dl.acm.org/citation.cfm?id=1927006.1927023>
- [50] M. Rodriguez, S. Ali, T. Kanade, Tracking in unstructured crowded scenes, in: 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 1389–1396.
url: <http://dx.doi.org/10.1109/ICCV.2009.5459301>
- [51] M. Ryoo, J. Aggarwal, Observe-and-explain: A new approach for multiple hypotheses tracking of humans and objects, in: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 2008, pp. 1–8.

- [52] P. Scovanner, M. Tappen, Learning pedestrian dynamics from the real world, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 381–388.
- [53] X. Shi, H. Ling, J. Xing, W. Hu, Multi-target tracking by rank-1 tensor approximation, in: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, 2013, pp. 2387–2394.
- [54] G. Shu, A. Dehghan, O. Oreifej, E. Hand, M. Shah, Part-based multiple-person tracking with partial occlusion handling, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 1815–1821.
- [55] F. Solera, S. Calderara, R. Cucchiara, Learning to divide and conquer for online multi-target tracking, CoRR abs/1509.03956.
url: <http://arxiv.org/abs/1509.03956>
- [56] B. Song, T.-Y. Jeng, E. Staudt, A. K. Roy-Chowdhury, A stochastic graph evolution framework for robust multi-target tracking, in: Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 605–619.
url: <http://dl.acm.org/citation.cfm?id=1886063.1886109>
- [57] D. Sugimura, K. Kitani, T. Okabe, Y. Sato, A. Sugimoto, Using individuality to track individuals: Clustering individual trajectories in crowds using local appearance and frequency trait, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 1467–1474.
- [58] V. Takala, M. Pietikainen, Multi-object tracking using color, texture and motion, in: Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, 2007, pp. 1–7.
- [59] S. Tang, M. Andriluka, B. Schiele, Detection and tracking of occluded people, Int. J. Comput. Vision 110 (1) (2014) 58–69.
url: <http://dx.doi.org/10.1007/s11263-013-0664-6>
- [60] A. F. Tchango, V. Thomas, O. Buffet, A. Dutech, F. Flacher, Tracking multiple interacting targets using a joint probabilistic data association filter, in: Information Fusion (FUSION), 2014 17th International Conference on, 2014, pp. 1–8.

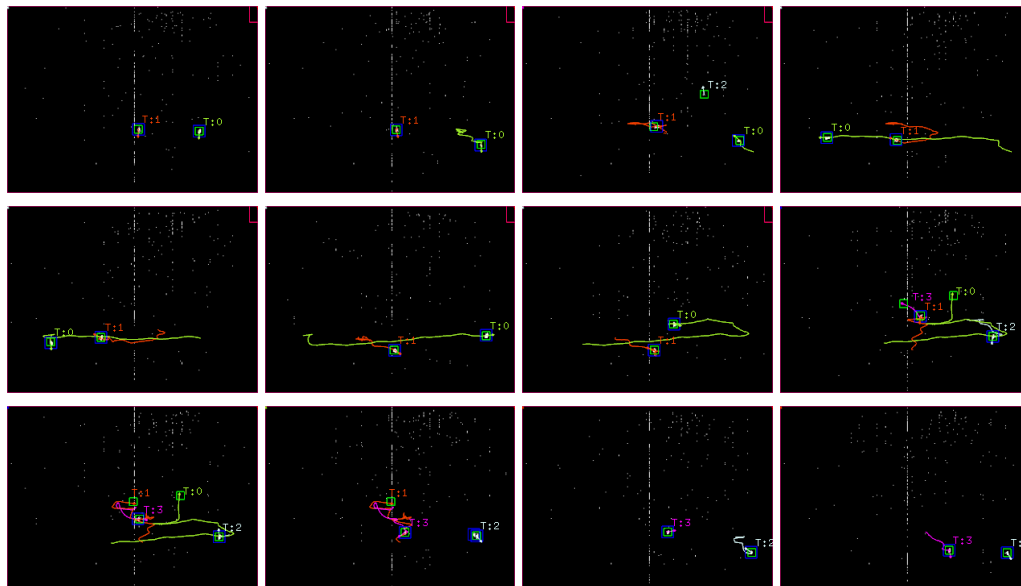
- [61] C. Tomasi, T. Kanade, Detection and tracking of point features, Tech. rep., International Journal of Computer Vision (1991).
- [62] T.-K. K. Wenhan Luo (Imperial College London), Generic object crowd tracking by multi-task learning, in: Proceedings of the British Machine Vision Conference, BMVA Press, 2013.
- [63] Z. Wu, A. Thangali, S. Sclaroff, M. Betke, Coupling detection and data association for multiple object tracking, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 1948–1955.
- [64] J. Xing, H. Ai, L. Liu, S. Lao, Multiple player tracking in sports video: A dual-mode two-way bayesian inference approach with progressive observation modeling, Image Processing, IEEE Transactions on 20 (6) (2011) 1652–1667.
- [65] J. Xing, H. Ai, S. Lao, Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 1200–1207.
- [66] K. Yamaguchi, A. Berg, L. Ortiz, T. Berg, Who are you with and where are you going?, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 2011, pp. 1345–1352.
- [67] B. Yang, R. Nevatia, Online learned discriminative part-based appearance models for multi-human tracking, in: Proceedings of the 12th European Conference on Computer Vision - Volume Part I, ECCV’12, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 484–498.
url: http://dx.doi.org/10.1007/978-3-642-33718-5_35
- [68] B. Yang, R. Nevatia, An online learned crf model for multi-target tracking, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 2034–2041.
- [69] M. Yang, F. Lv, W. Xu, Y. Gong, Detection driven adaptive multi-cue integration for multiple human tracking, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 1554–1561.

- [70] B. Yang, R. Nevatia, Multi-target tracking by online learning of non-linear motion patterns and robust appearance models, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 1918–1925.
- [71] A. Yilmaz, O. Javed, M. Shah, Object tracking: A survey, ACM Comput. Surv. 38 (4).
url: <http://doi.acm.org/10.1145/1177352.1177355>
- [72] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, L.-Q. Xu, Crowd analysis: A survey, Mach. Vision Appl. 19 (5-6) (2008) 345–357.
url: <http://dx.doi.org/10.1007/s00138-008-0132-4>
- [73] L. Zhang, Y. Li, R. Nevatia, Global data association for multi-object tracking using network flows, in: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 2008, pp. 1–8.
- [74] X. Zhao, D. Gong, G. Medioni, Tracking using motion patterns for very crowded scenes, in: Proceedings of the 12th European Conference on Computer Vision - Volume Part II, ECCV'12, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 315–328.
url: http://dx.doi.org/10.1007/978-3-642-33709-3_23

A Priloge



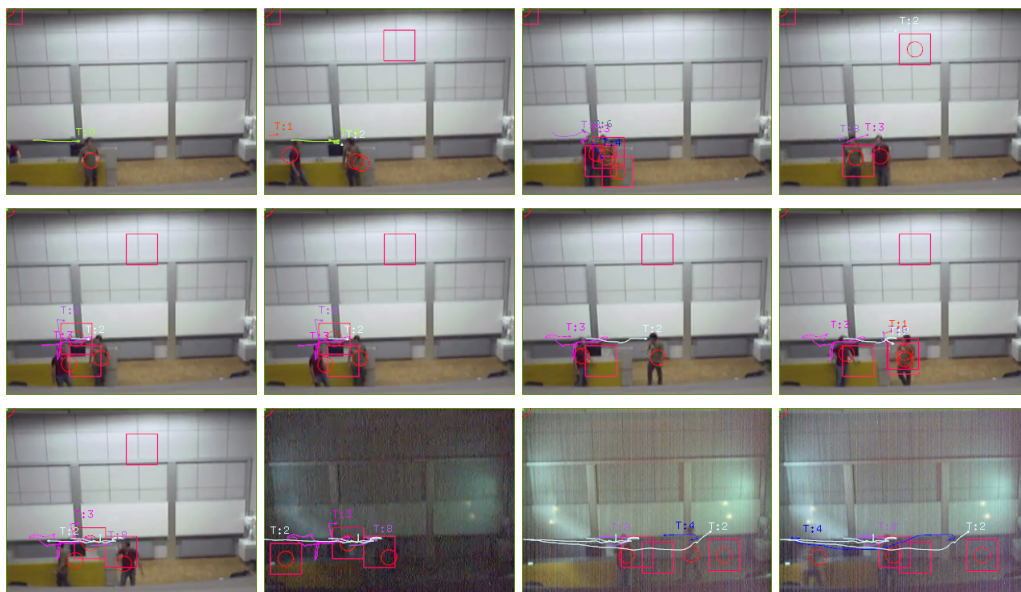
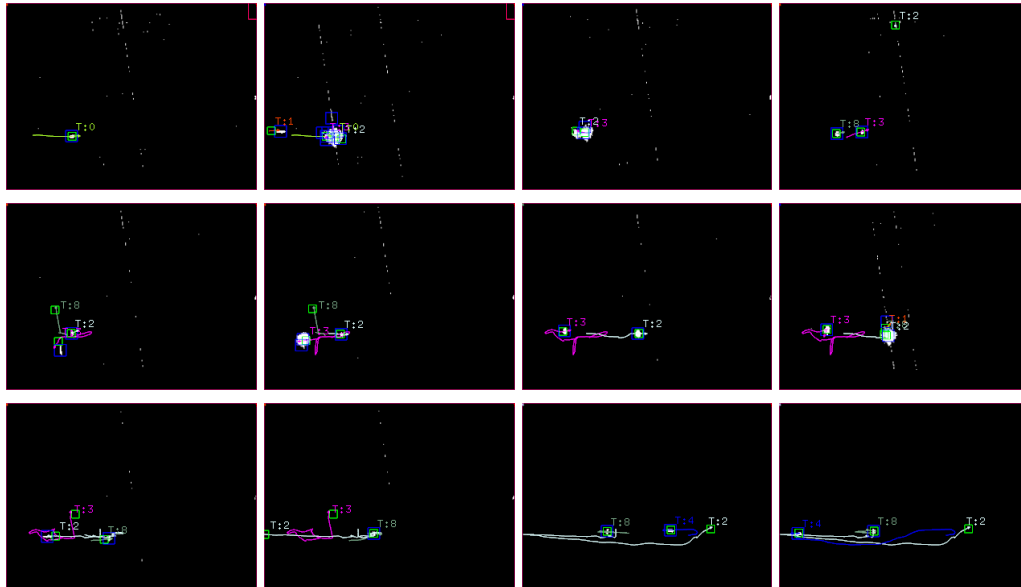
A.1 Slikovno zaporedje sledenja - 1. scenarij



Slika A.1 Rezultati sledilnika IR v spektru IR za 1. scenarij.

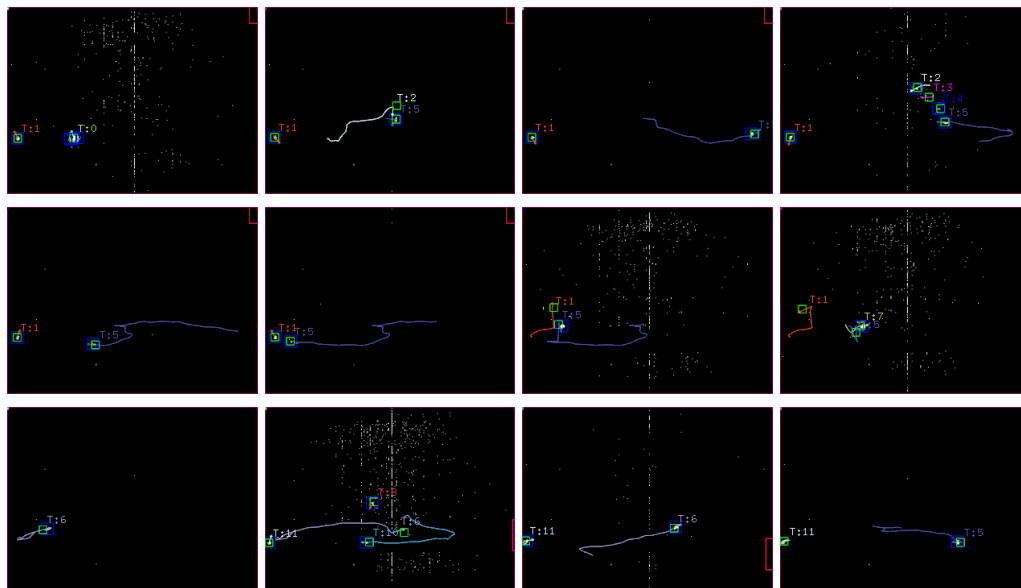


Slika A.2 Združeni rezultati sledenja obeh sledilnikov v vidnem spektru za 1. scenarij.

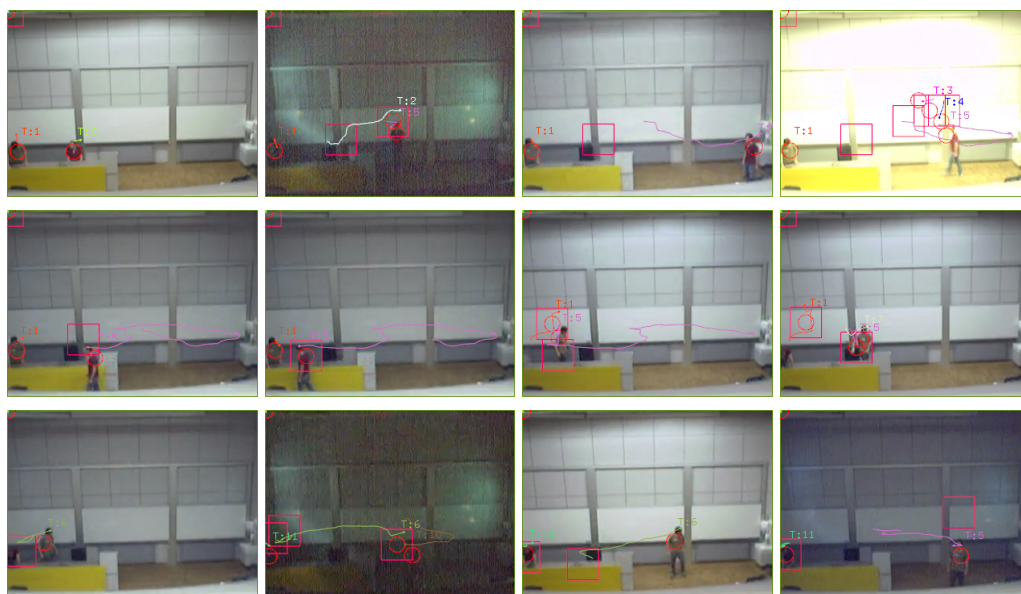


Slika A.4 Združeni rezultati sledenja obeh sledilnikov v vidnem spektru za 2. scenarij.

A.3 Slikovno zaporedje sledenja - 3. scenarij

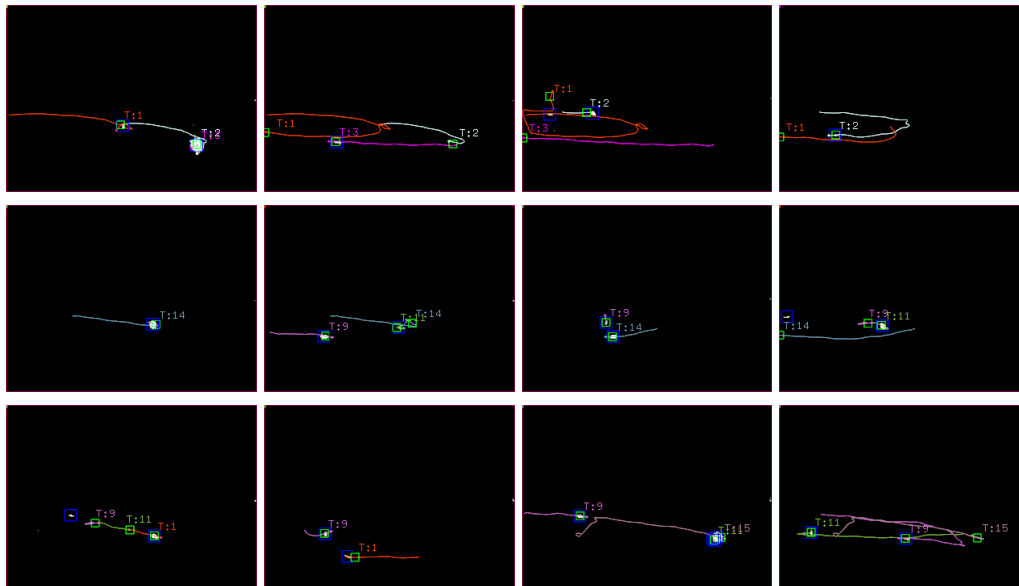


Slika A.5 Rezultati sledilnika IR v spektru IR za 3. scenarij.

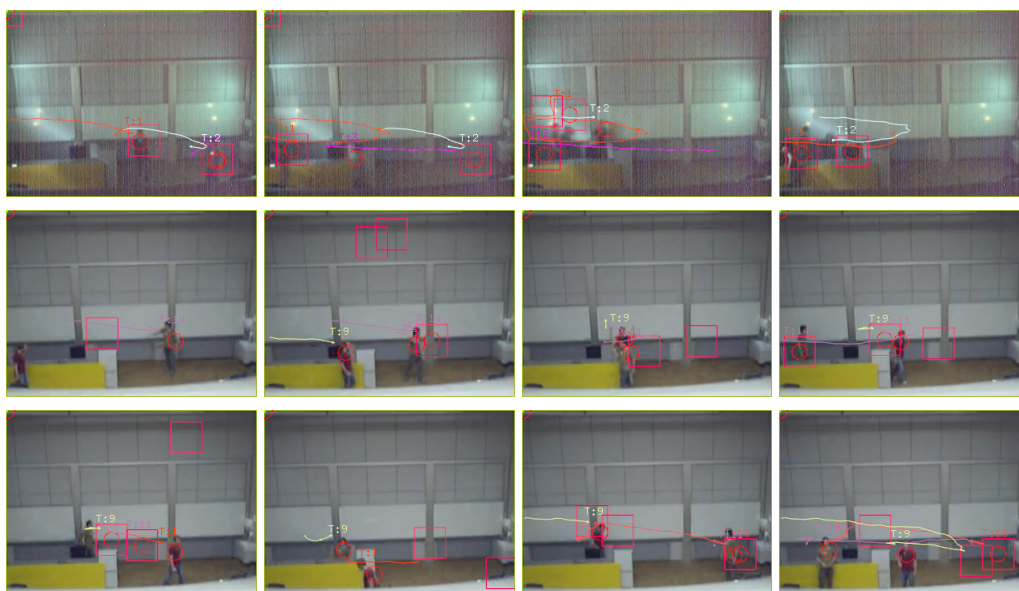


Slika A.6 Združeni rezultati sledenja obeh sledilnikov v vidnem spektru za 3. scenarij.

A.4 Slikovno zaporedje sledenja - 4. scenarij



Slika A.7 Rezultati sledilnika IR v spektru IR za 4. scenarij.



Slika A.8 Združeni rezultati sledenja obeh sledilnikov v vidnem spektru za 4. scenarij.